

**The Quality of Care for Adults with Mental and Addictive Disorders: Issues in
Performance Measurement**

Constance M. Horgan

Deborah W. Garnick

Revised -- January 10, 2005

Center for Behavioral Health

Schneider Institute for Health Policy

Heller School for Social Policy and Management

Brandeis University

Assistance from Elizabeth Merrick, Sharon Reif and Joanna Volpe-Vartanian as well as helpful comments from Ann Page are gratefully acknowledged. Partial support was provided by the Brandeis/Harvard NIDA Center on Drug Abuse and Managed Care (Grant # 2 P50 DA010233-06A2).

Background Paper prepared for the
Committee on Crossing the Quality Chasm: Adaptation to Mental Health and Addictive Disorders

The quality of behavioral health care can be measured, monitored, and improved over time using performance measures. In this paper, we first focus on describing the state of the art of performance measures for mental and substance use disorders. Next we explore the measurement challenges to the development of performance measures in behavioral health. Then we discuss the barriers to the widespread adoption and implementation of these measures. Finally, we propose systems changes that could accelerate the development and use of performance measures in behavioral health. Our discussion is based on the recognition that performance measures are tools and, as such, do not lead to improvements in quality of care unless they are well designed, appropriately used and applied in a system or organization that is equipped to implement change

The concept of performance measurement is well established in the medical sector where most purchasers, clinicians and policymakers, and increasingly some consumers now take for granted and indeed expect that performance measures for medical conditions will be calculated, published, and scrutinized. The development of performance measures in behavioral health has a more recent history; although performance measures for mental and addictive disorders are now receiving greater attention than ever before. The fact that this attention is warranted was highlighted by a recent national study showing that persons with depression received effective care only 57.7 percent of the time and that persons with alcohol dependence received effective care only 10.5 percent of the time, the lowest of any of the conditions examined (McGlynn et al., 2003). Great progress has been made in performance measures for behavioral health in recent years; nonetheless, some measurement issues and barriers to implementing widespread use of performance measures and quality improvement initiatives remain.

STATE OF THE ART IN BEHAVIORAL HEALTH PERFORMANCE MEASUREMENT

Historical Perspective

Challenges have arisen in the development of performance measures for mental health and substance abuse because of the structure of the treatment system, the quality and availability of data, and many complex computational issues. Measuring the quality of mental health and substance abuse services is particularly challenging for a variety of reasons. In mental health, no confirmatory laboratory or radiological test exists for diagnosis; clinical diagnosis is frequently

imprecise; important care processes often are not captured in data systems; clinical outcomes are not captured in a standard format; and performance is likely to be highly related to case mix. The situation is similar for substance use disorders; however biological tests are common for detecting the presence of a substance although this does not necessarily indicate a substance use disorder.

Historically, substance abuse and mental health performance was assessed through research studies that tracked clinical outcomes through post-treatment follow-up, typically at six months or one year after discharge (McLellan et al., 1998) or through analyses of administrative data on readmission within a short time period following discharge (Humphreys & Weingardt, 2000). Each approach had disadvantages. Clinical outcomes required costly follow-up interviews with clients who were in many cases difficult to locate. Administrative data on readmission rates focused on post-inpatient treatment time frames which were, by definition, short-term, typically within one week or one month.

Recent developments that go beyond those approaches emerge from several major initiatives that are described in more detail later in this paper. Among them are: the development of substance abuse performance measures by the Washington Circle (Garnick et al., 2002b; McCorry et al., 2000); the ongoing development and use of data collection standards and patient surveys through the Mental Health Statistics Improvement Project (Ganju et al., 2004; Teague et al., 1997); state level performance measure initiatives motivated, in part, by the measures being developed under the Substance Abuse Prevention and Treatment (SAPT) Block Grants (Curie et al., 2004); and clinically focused work such as the Patient Outcomes Research Team (PORT) on schizophrenia (Lehman et al., 2004).

Framework for Performance Measurement

In order to categorize and evaluate quality measures for behavioral healthcare systematically, it is key to begin with a framework. Referring to Donabedian's classic and oft cited work (Donabedian, 1980), quality measures can include structure, process and outcome measures as well as access and patients' experience (Agency for Healthcare Research and Quality, 2004). Each type of measure needs to be considered in terms of a specific use, however, because the assessments of cost, accuracy, aggregation and data source all depend on the purpose of a performance measure.

Structure. A structure measure is a feature of a healthcare organization, provider practice, or health system relevant to its capacity to provide health care. For example, the existence of an electronic medical records system or the ratio of mental health or substance abuse practitioners to clients are structure-based measures because they do not describe care given to specific patients or specific groups of patients.

Process. A process measure assesses a health care service provided to, or on behalf of, a patient. Process measures often are used to assess adherence to recommendations for clinical practice based on evidence or consensus, such as the extent to which clinicians adhere to guidelines for medication management for patients with depression.

To a greater extent than outcome measures, process measures can identify specific areas of care that may require improvement. Ideally, process measures are selected where scientific studies have established a link between the probability of achieving desired outcomes and the provision of particular services (McGlynn, 1998). Evidence for these links may come either from randomized controlled trials or from observational studies, such as one that showed that a positive relationship between guideline-concordant care and clinical depression outcomes in routine practice settings (Fortney et al., 2001).

Generally, a process measure is expressed as a rate with the denominator defining a population of interest based on its demographic and clinical condition and the numerator defining the subgroup receiving specific services. For substance abuse, for example, a process measure is the percent of adult enrollees in a managed care plan with a new diagnosis of substance abuse who receive timely follow-up services (Garnick et al., 2002b; McCorry et al., 2000). For mental health, among the many process measures for the treatment of schizophrenia (Hermann et al., 2002a), one example is the percent of adults hospitalized for an acute episode of schizophrenia who were prescribed an antipsychotic medication on discharge (Lehman et al., 2004; Lehman & Steinwachs, 1998). While process measures can be defined as simple rates, the technical specifications often are complex and need to include details about type of diagnoses or services, time periods, and data sources.

Outcome. An outcome measure is a health state of a patient resulting from health care practices and interventions. An outcome measure can be used to assess quality of care to the extent that health care services influence the likelihood of desired health outcomes.

Outcome-based measures of quality reflect the cumulative impact of multiple processes of care (AcademyHealth, 2004). Outcomes can be considered both in terms of categories of symptom severity related to the patient's clinical problem or in terms of functioning. For example, in substance abuse, outcomes can be considered in four domains: sustained reduction in alcohol and drug use, increases in personal health, sustained improvements in functioning (e.g., employment), and sustained reductions in threats to public health and safety (McLellan et al., forthcoming).

Access. An access measure assesses the patient's attainment of timely and appropriate health care. Barriers to access may include inability to pay for health care, difficulty traveling to health care facilities, unavailability of health care facilities, lack of a "medical home," cultural and health beliefs that prevent recognition of the need for and benefits of health care, and disparities in responding to persons seeking health care. This is particularly important in behavioral health because there is a large gap between those who have a diagnosed mental health or addictive disorders and those who actually use any services (Bijl et al., 2003; Regier et al., 1993a; Weisner & Schmidt, 2001).

An example of a substance abuse access measure is the identification measure adopted by the National Committee on Quality Assurance (NCQA). This measure is defined as the percentage of health plan members with alcohol or other drug abuse or dependence and a related service during the measurement year (NCQA, 2004a). An example of a mental health access measure is the percent of individuals who are offered an appointment within three business days among those requesting an initial appointment with a psychiatrist for medication management or assessment during a three month period (Hermann et al., 2002a). Another access measure is the mean duration between request for services by consumers with serious persistent mental illness and the first face-to-face visit (MHSIP Task Force, 1996; Network for the Improvement of Addiction Treatment (NIATx)).

Patient Experience. A patient experience measure aggregates reports of patients about their observations of and participation in health care, providing information on the patient perspective on quality of care. Frequently, when patient experience data is collected, other types of

information can be collected at the same time, e.g., measures of outcomes that can only be obtained from the patient such as reduction in symptoms, changes in patterns of substance use, or ability to work, attend school, and perform other activities of daily living.

In the behavioral health area, the Experience of Care and Health Outcomes Survey (ECHO™) is designed to collect consumer's ratings of their experience with specialty behavioral health care, including mental health, alcohol and drug, and other substance abuse services. Some key dimensions are getting care quickly, communication with clinicians, information provided by clinicians on medication side effects and family involvement in care (ECHO Development Team, 2002). The Mental Health Statistics Improvement Project (MHSIP) Consumer-Oriented Mental Health Report Card is an effort begun over 25 years ago to collect information from consumers (Ganju, 1996; Ganju et al., 2004; Teague et al., 1997).

In the medical sector, new evidence shows that high ratings in the areas of communication, interpersonal aspects of care and access are associated with better patient adherence to medical care. In behavioral health, the evidence is mixed. One study of addiction treatment shows that satisfaction is not a proxy for improvement in symptoms or functioning (McLellan & Hunkeler, 1998) although other work shows an association between satisfaction and the technical quality of care for anxiety and depressive disorder (Edlund et al., 2003).

Criteria for Measure Selection

Selecting performance measures is generally based on a set of criteria that includes the following three concepts: importance, scientific soundness, and feasibility. A variety of organizations have published criteria that fall broadly into these three (Agency for Healthcare Research and Quality, 2004; Hermann & Palmer, 2002; Joint Commission on Accreditation of Healthcare Organizations, 2004; McGlynn, 1998; NCQA, 1998).

Importance. This criterion encompasses the concepts of relevance to stakeholders; health importance in terms of high prevalence or incidence and a significant effect on the burden of illness to a population; potential for improvement; and susceptibility to being influenced by the health care system.

Scientific soundness. This concept can be considered in two areas. First, sound clinical logic means that the evidence supporting the measure is explicitly stated and this evidence strongly

supports the topic area of the measure. Second, sound measurement properties include reliability, validity, and allowance for adjustment for patient or consumer factors if required.

Feasibility. This criterion includes the requirement for explicit specification of the measure (generally expressed as a rate), data availability, the appropriateness of the cost or burden of measurement and the capacity to support subgroup analyses. Comprehensibility to the users of the measure also is key.

Along with these criteria, Hermann and Palmer (Hermann & Palmer, 2002) highlight specific dimensions that are particularly relevant for selecting measures in behavioral health, including selecting measures to focus on clinical populations, vulnerable groups, clinical setting, and level of health care system. In terms of clinical populations, for example, those considering a set of performance measures first need to decide between using general measures (e.g., time between first contact and appointment) that pertain for all patients and using measures that are focused on specific conditions. Next, they need to decide about whether to focus on conditions that are highly prevalent such as major depression and anxiety disorders or on other conditions that are less prevalent such as bipolar disorder, schizophrenia (Hermann et al., 2002a; Schizophrenia Patient Outcomes Team (PORT), 1998) and obsessive compulsive disorder but have a serious impact on functioning. Of course, it is key to recognize that some conditions, such as substance abuse, may be highly prevalent but under recognized in the data sets used for some performance measures (Lee et al., 2004).

While criteria are generally agreed upon, the relative weight placed on various dimensions by various stakeholder groups may differ. For example, one goal of the National Committee on Quality Assurance (NCQA) is to accredit health plans. Thus, it seeks measures appropriate for comparisons across plans and may place more weight on measures that differentiate among health plans. A behavioral health facility with the goal of instituting quality improvement initiatives within its own organization and examining effectiveness over time may place more weight on measures that are especially relevant for these initiatives. Often, a panel selecting measures for an organization might discuss alternative selection criteria, choose a list of criteria, collect candidate measures, and use a rating and scoring approach to narrow the list of measures before coming to consensus about the recommended set of measures (Organization for Economic Cooperation and Development, 2003).

In addition to selecting specific measures, often it is important to use a set of measures that represent the entire process of care (McCorry et al., 2000). Sometimes this goal is difficult to achieve because ideal measures may not be financially or technically feasible for some health plans or treatment systems to measure at the present time. Thus, as an intermediate approach, some groups have targeted a small number of process measures with the longer-range objective of expanding the measure set to include the entire process of care. This tactical approach of using a set of measures offers the opportunity to measure performance, albeit not the full process of care, while more sophisticated measurement approaches and more complete datasets are developed. However, this approach may lead managed care plans, policy makers, or clinicians to focus on the specific measure and not pay enough attention to key areas that are not measured. To circumvent this problem, sometimes measures are rotated.

Purpose of the Measure

Generally, the major purposes of performance measures include accountability, quality improvement and research (Solberg et al., 1997). More recently, performance measures are also used in “pay for performance” approaches in which providers or provider groups receive financial incentives based on their performance. Often, the same measures can serve multiple purposes depending on the setting where they are used, including state mental health or substance abuse systems, health plans or Medicaid programs, treatment facility, provider groups, or individual practitioner levels.

Accountability. Historically, accrediting bodies and regulatory agencies responsible for ensuring external accountability examined descriptive documentation of health care system structure and process. More recently, performance measurement has become a common means of ensuring external accountability (Horgan et al., 2004b). As the principal audience for this information, purchasers, payers, and health care consumers use performance measures to compare across different health care organizations so the focus is on common definitions and comparability of information. Additionally, managed care plans and state agencies that contract with managed behavioral healthcare organizations use performance measures to hold their contractors accountable for the quality of service they provide (Garnick et al., 2001).

Quality improvement. Quality improvement can be considered as an activity internal to an organization in which the main objective often is to obtain results in a timely manner that can be

used to drive quality improvement. For example, substance abuse treatment agencies that are members of the Substance Abuse and Mental Health Services Administration (SAMHSA) and Robert Wood Johnson Foundation-funded Network for the Improvement of Addiction Treatment (NIATx) have selected areas of performance that need improvement (e.g., reduce waiting time between first request for service and first treatment session), measured the extent of the problem, developed strategies for improvement, and repeated the measurement. Many organizations use continuous quality improvement (CQI) approaches as part of internal monitoring and management.

CQI also can be externally imposed. Performance measures are used by organizations to drive quality improvement by influencing the actions of others. For example, health plans can write performance standards into their contracts with managed behavioral healthcare organizations (MBHOs) where specific performance goals must be achieved with stipulation of specific targets for improvement over time (Horgan et al., 2004b). State agencies can do profiling and feedback to provider organizations (Oklahoma Department of Mental Health and Substance Abuse Services, 2003) or can influence quality by basing payment on performance (Shapiro, 2004).

Research. Often performance measures are used in the context of policy research. For example, investigators evaluating the impacts of a change in policy, such as increased cost sharing for behavioral health services, could add quality measures to their set of predicted effects along with costs and utilization.

Overview of Current Health Performance Measures

Numerous measures have been developed to assess the quality of behavioral health care. Two major computerized inventories include behavioral health performance measures and offer opportunities for users to assess their suitability for specific applications by including information on the detailed specifications of the measures, the clinical evidence behind their development, and sources for further information.

First, the National Quality Measures Clearinghouse (NQMC), operated by the federal Agency for Healthcare Research and Quality, includes measures for medical, mental health and substance abuse conditions (Agency for Healthcare Research and Quality, 2004). Searching the NQMC for the mental disorders category in January 2005 yielded 64 measures, substance-related disorders

yielded 55 measures (including measures focused on tobacco), and alcohol-related disorders yielded 38 measures. The NQMC database is continuously updated and measures must meet six inclusion criteria: 1) address an aspect of care for a defined group of individuals, 2) relate to at least one quality domain (e.g., structure, process), 3) have documented supporting evidence, 4) have evidence of reliability and validity or endorsement by an organization that promotes rigorous development, 5) have evidence of current use, and 6) include detailed English-language documentation.

Second, the National Inventory of Mental Health Quality Measures includes approximately 300 measures for both mental health and substance abuse (Hermann et al., 2002b). Unlike the NQMC which allows searching by the broad array of terms listed under the National Library of Medicine's medical subject headings (MeSH), the National Inventory of Mental Health Quality Measures offers menus of search parameters including diagnoses (e.g. depression, personality disorders, substance abuse/dependence), population, clinical setting, treatment approach (e.g., medication, psychotherapy), data source, and evidence level. Criteria for measure inclusion are not listed on the website.

Table 1 shows examples of mental health and substance abuse performance measures, all of which can be found in these two inventories. For each measure, we show the domain, a brief definition of the measure, the target population, the data source and reference for where to obtain more detailed information. In selecting measures to show as examples, our goals were to present examples in all five domains (structure, process, outcome, access and patient experience), as well as to offer a range of data sources, target populations, and measure developers. The examples we selected range from the substance abuse measures newly adopted by the National Committee on Quality Assurance (NCQA) for commercial health plans, Medicaid and Medicare managed care to a measure of enrollee satisfaction used by a single state Medicaid program.

Table 1. Selected Behavioral Health Performance Measures

Domain	Measure	Target Population	Data Source
Structure	<i>Provider’s mean score on Competency Assessment Instrument (CAI):</i> measures 15 competencies needed to provide high quality care for severe and persistent mental illness and can be scored as a summary or separate scales (e.g., community resources, medication management) (Chinman et al., 2003)	Providers of services to those with serious mental illness	Clinician survey
Process	<i>Initiation of Alcohol and Other Drug (AOD) Treatment:</i> Percentage of adults diagnosed with AOD who initiate treatment through one of the following: <ul style="list-style-type: none"> • an inpatient AOD admission, <i>or</i> • an outpatient service for AOD abuse or dependence <i>and</i> any additional AOD services within 14 days. <i>Treatment engagement:</i> Percentage of adults who engage in treatment with two additional AOD treatments within 30 days after initiating treatment (Garnick et al., 2002b; NCQA, 2004a)	Adult health plan members with AOD diagnosis and related service	Admin. data
	<i>Major depressive disorder:</i> Percentage of patients aged 18 years and older with confirmed diagnosis of major depressive disorder who had a suicide risk assessment completed at each visit (Physician Consortium for Performance Improvement)	Patients 18 and older with diagnosis of major depressive disorder	Medical record
	<i>Mental illness:</i> Percentage of members who had an ambulatory or day/night mental health visit within 30 days of hospital discharge (NCQA, 2004a)	Members hospitalized for selected mental health disorders	Admin. data
	<i>Antidepressant medication management (optimal practitioner contacts for medication management):</i> Percentage of members who had three or more outpatient follow-up visits or day/night treatment within (12 weeks) after new diagnosis of depression (NCQA, 2004a)	Patients 18 and older with diagnosis of major depressive disorder	Medical record
Outcome	<i>Major depression in adults:</i> Percentage of patients whose results on 2 quantitative symptom assessment tools (such as Patient Health Questionnaire [PHQ-9]) decrease by 50 percent within six months of initiating treatment (Institute for Clinical Systems Improvement (ICSI), 2004)	Primary care patients 18 and older with new diagnosis of major depression	Patient survey
Access	<i>Identification of Alcohol and other Drug (AOD) Services:</i> percentage of members with an alcohol and other drug (AOD) claim. (Garnick et al., 2002b; NCQA, 2004a)	Adolescent and adult health plan members	Admin. data
	<i>Major depressive disorder:</i> Percent of patients screened for depression (excludes those in treatment for depression) (Veterans Health Administration, 2002)	Individuals seen at least once in primary care during a 12-month period	Admin. data and medical records data
Patient Experience	<i>Satisfaction with referral for mental health/substance abuse care:</i> Percentage of enrollees who reported getting a referral (Wisconsin Department of Health and Family Services, 2003)	Wisconsin Medicaid/Badger Care enrollees	Patient survey

0000Many groups have been involved in developing measures or in endorsing and disseminating measures that others have developed including:

- ***Federal agencies*** such as the Department of Veterans Affairs, Substance Abuse and Mental Health Services Administration (SAMHSA), and Agency for Healthcare Research and Quality (AHRQ)
- ***Professional clinician associations*** such as American Psychiatric Association and the American Association of Child and Adolescent Psychiatry
- ***Independent organizations with a behavioral health focus*** such as the Washington Circle Group which is focused on substance abuse, the American Managed Behavioral Healthcare Association (AMBHA), and the Forum on Performance Measurement in Behavioral Health and Related Service Systems
- ***Independent organizations with a general focus*** such as the National Committee on Quality Assurance (NCQA), Joint Commission on the Accreditation of Healthcare Organizations (JCAHO), National Quality Forum (NQF), or the Commission on Accreditation of Rehabilitation Facilities (CARF)

Through the websites that catalog measures or the websites of these developers, it is now possible for users to identify many mental health and substance abuse performance measures.

Examination of these sources reveals, however, that often multiple measures focus on the same concept or on measures that are similar, but not exactly the same, an issue we discuss later in this paper.

Effectiveness of Treatment

Over the past decade, research in the fields of mental health and substance abuse has offered consistent scientific evidence that some specific practices work well in improving outcomes in the lives of individuals diagnosed with mental illness or substance disorder. The role of this body of research, as it relates to performance measurement, is to provide the scientific basis upon which to build measures of performance at the provider, practice, or organization or system levels.

Determining the effectiveness of many types of treatment for behavioral health is sometimes more challenging than for medical care. Consider the medical example of using beta-blocker medications to treat heart disease where treatment is the timely administration of the right dose of medication and outcomes relate to cardiac-related morbidity and mortality. This reflects a specific activity that is relatively straightforward to measure both from the perspective of process of care and patient outcomes. Some behavioral health process and outcome measures may be comparable, such as studies of buprenorphine for the treatment of opioid addiction. Other behavioral health treatment approaches, however, present more complex challenges because of the need to measure treatments that involve multiple elements of counseling, education, structured interviewing or other approaches that may be more difficult to standardize or quantify and the need to measure outcomes that may include social functioning and reduction in alcohol or drug use.

Use of clinical trials. In considering evidence for treatment effectiveness, two questions often arise: first, whether clinical trials are necessary to establish effectiveness; and second, whether evidence from clinical trials is generalizable to ordinary practice settings. When a consensus on treatment effectiveness is sought, observational studies sometimes are included along with controlled trials with random assignment or evaluations and demonstrations, but studies often are weighted based on a set of criteria that relate to scientific method (Miller & Wilbourne, 2002). Humphreys and Weisner (Humphreys & Weisner, 2000) point out that exclusion criteria in clinical trials of alcohol treatment can result in research samples that are more heavily composed of white, economically stable, and higher-functioning individuals than are real-world samples of substance abuse patients, potentially compromising the generalizability of results.

Concern also has been expressed that randomized clinical trials may be problematic in evaluating socially complex interventions that may be common with individuals who have mental or addictive disorders. In particular, the environment may interact with the intervention complicating interpretation of results if these contextual variables are not controlled for (Wolff, 2001). For example, an individual may receive a particular clinical intervention which might be complicated if the person becomes homeless.

New practices research and approaches. Because new practices are continuously being developed and new research is ongoing about effectiveness, conclusions regarding effective practices also require continual updating. In addition to new treatment approaches, new ways of conceptualizing effectiveness also have been recently proposed. In particular, McLellan and colleagues (McLellan et al., forthcoming) propose a new evaluation system for outpatient addiction treatment called *concurrent recovery monitoring*, in which outcome measures are collected immediately on beginning treatment and continually through out the course of treatment. This new approach differs from previous outcome evaluations of addiction treatment that have followed patients at one or two discrete times following discharge.

Because evidence on treatment effectiveness is drawn from hundreds of studies, it is useful to turn to summaries such as the Mesa Grande project (Miller & Wilbourne, 2002) that relates to alcohol use disorders, the Cochrane Collaboration that relates to treatment for drug and alcohol addiction and mental health (Cochrane Depression Anxiety and Neurosis Group, 2004; Cochrane Drugs and Alcohol Group, 2004), the evidence compiled by the World Health Organization on the effectiveness of treatment for mental health problems in the elderly (World Health Organization Regional Office for Europe, 2004), and the Treatment Improvement Protocols (TIPs) published by the Center for Substance Abuse Treatment (CSAT) of the Substance Abuse and Mental Health Services Administration (SAMHSA, 2004d).

Of course, studies of treatment for alcohol, drug and mental health problems may reveal practices that are not effective, although these practices may remain widespread. In this section, therefore, we provide examples both of treatment that the published literature shows to be effective and other examples where the evidence points to the opposite conclusion. It is key to note that these are offered as examples only, selected to show a range of mental health and substance abuse treatment practices.

Examples of Effective Practices

Psychotherapies and medication for depression. An accumulating body of evidence demonstrates that both time-limited psychotherapies and antidepressant treatments are effective in psychiatric and primary care settings (Schulberg et al., 1998). Generally, either medication combined with four or more follow up visits for monitoring, or a minimum of eight visits with a mental health specialist in the absence of medication, conforms with evidence-based guidelines

for depression care during the acute and continuation phases of treatment (Agency for Healthcare Research and Quality, 2004; American Psychiatric Association, 1993; Wang et al., 2000; Wang et al., 2002).

Time in treatment for drug and alcohol disorders. Studies of treatment retention indicate that patients who remain in treatment for a longer duration of time have better outcomes (Grella et al., 1999; Hubbard et al., 2003). This positive association between retention and treatment success, along with evidence about treatment intensity (Fiorentine & Anglin, 1996; McLellan et al., 1993; Simpson et al., 1995), provided the rationale for the substance abuse initiation and engagement measures developed by the Washington Circle and adopted by the NCQA. In brief, becoming engaged in treatment with an initial intensity of services was considered to be an important intermediate process measure that is closely related to outcomes (McCorry et al., 2000).

Brief interventions in primary care for alcohol-related problems. Primary care providers commonly treat patients who are experiencing a range of alcohol-related problems and risks (Whitlock et al., 2004). As a result, these clinicians are optimally positioned to identify alcohol problems in their patients, an essential first step, and to intervene (Friedmann et al., 2000). Evidence demonstrates that screening and brief intervention methods in primary care have been found to markedly reduce drinking levels in patients (Fleming et al., 1999; Wilk et al., 1997). Based on this evidence and their own research, Babor and Higgins-Biddle recommended a survey based performance measure of querying patients as to whether their physicians had asked them about drinking (Babor & Higgins-Biddle, 2003).

Examples of Ineffective Practices

Confrontational methods for treatment of alcoholism. These methods, including video self-confrontation, confrontational counseling, and educational lectures and films were all shown to be ineffective showing negative “cumulative evidence scores” according to the rating methods devised by Miller and Wilbourne (Miller & Wilbourne, 2002).

Antidepressants in treatment of cocaine dependence. There is no current evidence supporting the clinical use of antidepressants in the treatment of cocaine dependence according to a review of eighteen studies by the Cochrane Collaborative. Moreover, similar results were obtained for trials where patients had additional diagnosis of opioid dependence and/or were in methadone maintenance treatment (Lima et al., 2004).

Bringing Effective Practices to the Field

Knowledge of the evidence of effective practices may not quickly reach the front lines of treatment, in part because of the rapid proliferation of new information, proposals for new treatment based on new understanding about the chronic nature of some mental health and addiction problems, and the lack of a well developed infrastructure which would support diffusion of practices. Performance measures can be developed, in concert with improved infrastructure, so that information on both effective and ineffective interventions can reach practitioners more quickly.

Evolution of Therapies and Medication. Extraordinary progress in our knowledge of the brain, behavior, and effective treatments for many addictive and mental disorders has occurred in recent years. Evidence-based medications for specific conditions and psychosocial interventions now permit the successful treatment of many of these conditions. In many cases, the most effective treatment involves the combination of psychopharmacologic treatment and psychosocial interventions. Promising but less thoroughly documented treatments and services, sometimes called emerging best practices, are also available (New Freedom Commission on Mental Health, 2003). Addictive disorders have had a more recent history of the employment of medications in treatment; and use of medications, such as naltrexone for alcoholism and buprenorphine for opioid addiction are earlier in the diffusion process (NIAAA, 2000; NIDA, 2004).

Despite the productivity of scientific research, there is remarkable consensus that both evidence-based and emerging best practices for addictive and mental disorders are slow to be incorporated into service delivery (Institute of Medicine, 1998; New Freedom Commission on Mental Health, 2003; NIDA, 2004). The IOM report (2001) on the quality of American health care has noted the pervasiveness of this lag problem in the medical area with diffusion into routine care taking from 15 to 20 years (Balas & Boren, 2000). There is no evidence to suggest that the lag time is any shorter for addictive and mental disorders. Even when treatment approaches are routinely used in practice, there is concern about fidelity to the original treatment model resulting in clinical practice that may be inconsistent (Drake et al., 2001).

Process of Care – Recovery and Maintenance. Many mental and addictive disorders are viewed for some individuals as having chronic aspects where treatment effects may be optimized by

monitoring strategies and continuing care, similar to the treatment of diabetes, asthma, and hypertension (McLellan et al., 2000) (US Department of Health and Human Services, 1999). This view is consistent with the concept of recovery introduced by consumers with mental illness in the 1980s in which the overarching theme is that hope and restoration of a meaningful life is possible, despite serious mental illness (Anthony, 1993). It is important to note that not all cases of mental and addictive disorders are chronic; however, for many individuals continuation of sustained improvements in health and social functioning may require efforts that focus on maintenance. This approach is consistent with the well-known chronic care model widely accepted in the medical arena for following patients with chronic illnesses (Bodenheimer et al., 2002; Wagner et al., 1996). This model is based on the premise that the conditions are continuing, rather than acute and service provision is adjusted to the severity of the symptoms at a given time. The applicability of this model to addictions (McLellan et al., forthcoming) and to mental illness has received increased attention. The chronic care approach necessitates a focus on services needed for the long-term management of both mental health and substance use disorders. To date much of performance measurement has focused on the front end of service delivery in terms of the process of care. Further development of performance measures will be necessary to capture the maintenance emphasis of the chronic care model.

Ongoing efforts. Several efforts are underway to help to ensure that treatment programs that are shown to demonstrate effectiveness are adhering to specific treatment practices. These efforts come in response to concerns about the fidelity of incorporating specific treatment practices into the real world. The National Registry of Effective Programs and Practices (NREPP) is a SAMHSA program that identifies and evaluates programs according to rigorous standards of research, and makes those model programs available to practitioners along with the materials, training, and technical assistance needed for implementation (SAMHSA, 2004b). The American College of Mental Health Administration (ACHMA) also has produced a manual with the same goals of educating practitioners about effective practices and giving them information on how to incorporate those practices into their treatment settings (AHCMA, 2003).

CHALLENGES IN MEASURING QUALITY OF MENTAL HEALTH AND SUBSTANCE ABUSE CARE

Beyond the need to better link evidence-based practice to real world settings, there are additional challenges to measuring the quality of mental health and substance abuse services. These include the nature of the service delivery system, types and availability of data, population issues, data quality and calculation issues.

Nature of Service Delivery System

The current system for mental health and addiction services remains fragmented across many dimensions including the simultaneous needs of many patients for services provided outside the health care system, the lack of integration between specialty behavioral health and primary care, and the separate treatment systems for mental health and substance abuse. Not only do these structural aspects of organization and financing challenge the actual delivery of high quality treatment services, they also compromises the ability to even measure whether the services are being provided. Greater integration across these various systems of care, both within the health care system and extending to the broader human service systems, and the ability to track service delivery in these multiple environments will be necessary as better performance and improved outcomes are demanded by stakeholders in the behavioral health area.

Multiple Systems. Multiple pathways to treatment characterize the nature of the service delivery system itself. Moreover, the organization and financing of services for these conditions is particularly complex because identification and treatment can occur in diverse health and human service settings (Weisner & Schmidt, 1993). These settings include the providers and entities that specialize in the treatment of mental and addictive disorders, while the general medical component includes primary care providers, emergency departments, and medical care specialists. The additional human services component consists of a broad array of social services related to the educational, welfare, and criminal justice systems, as well as the workforce. For example, most treatment of drug-involved offenders takes place in community-based settings, frequently as a condition imposed by probation or parole (Institute of Medicine, 1996). Employee assistance programs (EAPs), while primarily designed to address work performance issues, are another behavioral health resource that may be involved in direct service provision, especially related to identification and referral to treatment (French et al., 1997; Merrick et al., 2003; Roman, 2002). The voluntary support sector is not typically viewed as formal treatment; however, self-help

groups, such as Alcoholics Anonymous (AA) and other types of peer counseling play an important role in the recovery process for many individuals with mental and addictive disorders.

The relative importance and intersection of these various sectors in service provision is important to note because of the added complexities that are introduced in attempting to measure the quality of treatment when multiple sectors are involved in service provision. For example, epidemiological data has shown that among individuals with a diagnosable alcohol disorder, 22 percent received some mental health or substance abuse services in one or more of these types of settings over a twelve month period with 11 percent receiving care in the specialty sector, 10 percent in the medical sector, 8 percent through voluntary support networks, and 4 percent in the human services sector. The sectors are different in terms of types of data available for performance measurement and the linkability across subsystems. Indeed, the voluntary support sector is characterized by a lack of data and in some cases, a commitment to anonymity. Also, of note with respect to measuring performance, particularly related to access and problem identification are the 78 percent of individuals who receive no treatment (Regier et al., 1993b).

Some of the quandaries in ensuring quality and measuring performance in such a fragmented system are described below.

Primary vs. Specialty Care. Many individuals with mental health, alcohol and drug problems are encountered in the primary care or other specialized medical setting, with about half of all episodes of mental health and substance abuse care occurring in primary care (Institute of Medicine, 1996). It is estimated that in a given year, 10 to 20 percent of the population see a primary care provider about a mental health problem (Institute of Medicine, 1996). Mental and addictive disorders frequently co-occur with medical conditions (Saitz et al., 1997; Samet et al., 1996). Although a person may present in primary care for a physical condition, frequently there is an underlying or accompanying mental or addictive condition, which may or may not be recognized. When recognized, the treatment for the mental and/or addictive disorder may be subsumed under another diagnosis, particularly when it is seen as a secondary condition (Horgan & Merrick, 2001). Many studies have described the lack of recognition of mental health and substance abuse problems in the primary care setting (Institute of Medicine, 1997). The primary care setting is particularly important from the perspective of screening, brief interventions, referral to the specialty setting, as well as the treatment of other medical conditions concurrently

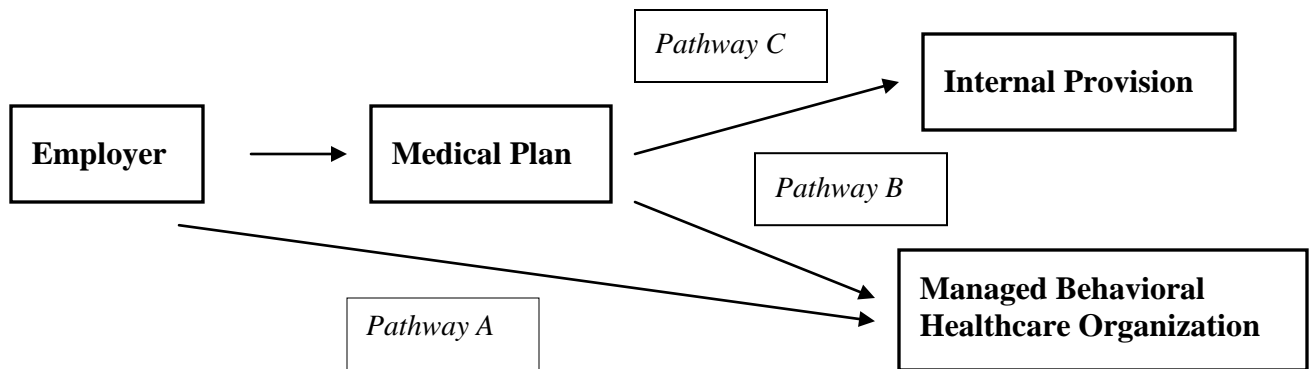
with mental, alcohol and drug problems (Friedmann et al., 2001; Samet et al., 1996). A quality concern particularly related to service provision in the primary care setting is the under recognition of behavioral health problems. It is a measurement challenge to identify a problem that is not appearing in any kind of data, suggesting the need to develop some kind of benchmarking so as to compare the identified population with what might normally be expected. Another concern relates to the quality of the behavioral care that is provided in a primary care setting, particularly for individuals with more severe conditions (Institute of Medicine, 1997). It is a measurement challenge to appropriately risk adjust across settings. For example, the outcomes of treatment in specialty and primary care settings for individuals with severe depression were compared, and the quality of care provided by psychiatrists was found to be significantly better than the quality for primary care physicians (Sturm & Wells, 1995). It is important to control for the level of severity of the client when measuring the performance of different types of providers. In addition to the issue of improving the quality of the treatment of mental and addictive disorders within the primary care setting, is the issue of the need for better coordination between both settings. Indeed the recent presidential commission on mental health described a “chasm” between mental health care and general health care systems with respect to both financing and practice (New Freedom Commission on Mental Health, 2003).

Managed Behavioral Health Care. Specialty mental health and substance abuse services are often combined into the broader area of managed behavioral health care with services frequently provided through specialized managed behavioral health organizations (MBHO). Managed behavioral health care for the privately insured, as well as those covered by Medicaid and other public funding, has grown rapidly through the 1990s and into this decade (Horgan & Merrick, 2001).

Figure 1 shows two alternative pathways in contracting with MBHOs. Purchasers, such as employers or state Medicaid agencies, can contract directly with an MBHO eliminating any connection with a health plan for specialty services (pathway A). Alternatively, purchasers can contract directly with a health plan that assumes responsibility for the provision of both medical and behavioral health services. The health plan itself may then choose to contract behavioral health services to an MBHO (pathway B) or to provide services internally (Pathway C) (Hodgkin et al., 2000; Horgan et al., 2000). Private health plans have been increasingly choosing pathway

B and contracting out the provision of specialty services to MBHOs with 58 percent contracting out in 1999 (Horgan et al., 2003) and 71 percent doing so in 2003 (Horgan et al., 2004b), although this trend may have reached its peak with ongoing discussion in some health plans to bring behavioral health back in house (Kazel, 2004). These specialty carve-out arrangements have important implications for performance measurement. In pathway A since the contract is directly with the purchaser, the data system would only be able to capture services provided by the MBHO and would not be able to link with health plan data, thus missing any services that might be provided through the health plan, including primary care and pharmacy data. In pathway B, since the contract is with the health plan, the MBHO data are in theory linkable with other health plan data, thus providing a greater opportunity for capturing the full array of behavioral health services and pharmacy claims.

Figure 1. Alternative Contracting Arrangements in Behavioral Health (adapted from (Hodgkin et al., 1997))



Separate Mental Health and Substance Abuse Systems. The specialty mental health and the alcohol and drug service systems have developed relatively independently of each other, and of the medical system as a whole reflecting a historical ambivalence to recognizing mental and addictive disorders as medical illnesses (Institute of Medicine, 1997; New Freedom Commission on Mental Health, 2003; US Department of Health and Human Services, 1999). Although

individuals with substance use disorders are treated in the specialty mental health system, and conversely individuals with serious mental illness are treated in the specialty alcohol and drug treatment system, concern has been noted that individuals with these conditions are less than adequately recognized and treated (Gerstein & Harwood, 1990; Institute of Medicine, 1990a). The mental health and substance abuse treatment systems are largely distinct with differences in patient populations, organizational workforce, treatment philosophies, as well as funding mechanisms (Horgan & Levine, 1998; Horgan 1997; Horgan & Hodgkin, 1999; Institute of Medicine, 1997). These distinctions present challenges when attempting to develop common performance standards and indeed in the ability to actually measure performance given underlying differences in data systems because of the need to consider individuals who have co-occurring mental health and substance use disorders.

Mix of Public and Private Funding. Within the specialty mental health and alcohol and drug treatment systems, a mix of federal, state, and private financing mechanisms target different treatment populations creating similar yet distinct service mixes (Institute of Medicine, 1990b; McCarty & Goldman, 2004). This array of funding mechanisms contributes to a system that is inherently fragmented which complicates the provision of quality of care to an individual whose needed services may cut across multiple funders and providers. More than half of funding for mental health and substance abuse treatment comes from public-sector payers. Approximately 62 percent of expenditures for substance abuse treatment and 55 percent of expenditures for mental health in 1997 came from the public sector including Medicaid, block grant funding, and other federal, state and local funding (Mark et al., 2000). Multiple mixes of services, providers and funding mechanisms present challenges to measuring performance. There is the concern that one may be only measuring a piece of the mix, thus being careful to define what part and what the implications are become key.

Access to Treatment

Both the entry into treatment and the receipt of services in a timely manner can influence the quality of mental health and substance abuse services, as discussed below.

Treatment entry. Facilitating access to behavioral health treatment is crucial because mental illness and addictive disorders, while common and often serious, are frequently not addressed

despite the availability of effective treatments. While not all those with these disorders seek treatment, it is important to look at reasons why not. It is particularly relevant to focus on system-level factors such as the way that entry into treatment is structured because it can affect access and is something that is under the control of the treatment system, and is therefore able to change. Furthermore, access is multi-dimensional, including initial contact with various types of care as well as continued services when needed, and is related to social factors such as definition of need. There are many factors affecting individuals' involvement in treatment beginning with initial entry (Weisner & Matzger, 2002; Weisner et al., 2002; Weisner & Schmidt, 2001) A 1998 national survey found that 11 percent of the population perceived a need for mental or addictive services and about 25 percent reported difficulties in obtaining care with worry about costs reported as the leading reason for not receiving care (Sturm & Sherbourne, 2001).

Timeliness. An essential dimension of access to behavioral health treatment is provision of timely initial appointments once individuals request care. A study reporting on a nationally representative survey of managed care organizations regarding how mental health services were provided in 1999 found that nearly half (43 percent) of all managed behavioral healthcare products reported having formal standards for maximum wait time from request for treatment to initial appointment (Horgan et al., 2003). In theory, quality standards should ensure reduced waiting lists and shorter delays for treatment. In practice, problems persist. Matevia and colleagues (Matevia et al., 2001) explored the issue of potentially declining access to network clinicians at a large MBHO and found lengthening waiting lists for both psychiatrists and non medical mental health specialists. Rost (Rost et al., 1994a) investigated physician preferences in treating depression in rural patients and found that most physicians reported recently referring at least one patient to specialty care and confronting limited referral sources and long waiting lists.

Types and Availability of Data

Data for measuring the quality of behavioral healthcare can come from multiple sources including claims and encounter data, medical records, and patient surveys. Each source has strengths and limitations in terms of cost, accessibility, timeliness, and accuracy. Moreover, each source should be matched to the particular goals of accountability, quality improvement, research, or pay for performance as discussed above (Garnick et al., 2002a).

Administrative (claims, encounter, eligibility). These individual-level data on utilization of and

charges for services are generated by insurers and related entities in the course of providing and paying for care. Often this source includes linkable information on pharmacy claims, which is particularly useful for tracking the quality of care for depression and other conditions in which medication management plays a key role.

Administrative data allow identification of patients receiving mental health, drug or alcohol or drug related services and tracking their care over time and across settings. Even without full standardization of data collection, certain key elements are usually present in claims data. These include amount billed and amount reimbursed (often separately by service), type of service, how many units (e.g. days), where provided, and diagnosis and procedure codes for clinical services.

For mental health, alcohol or drug treatment, even more than medical care, an important proportion of treatment is not reported on traditional insurance claims because the patient lacks insurance, is self-paying, or there is no requirement for the filing of claims for the provider to be paid. The last situation occurs in systems such as staff-model HMOs, state-run specialty hospitals and Veterans' Administration (VA) facilities, all of which deliver care through facilities they own and providers they employ. Most of these systems generate encounter-level records that share the common feature with claims of having dates and descriptions of services, although generally not information on payment.

Claims data are used for measuring quality of behavioral health care in a variety of settings including commercial health plans, Medicare, Medicaid and the Veterans Administration. For commercial health plans, as well as the Medicare and Medicaid programs, both the mental health and substance abuse measures in the Health Employer Data and Information Set (HEDIS) developed by NCQA and described above rely on administrative data.

In the VA, administrative data has been used to focus on quality measurement for mental illness in studies of antidepressant dosage (Charbonneau et al., 2004), compliance with dosing recommendation for schizophrenia (Leslie & Rosenheck, 2001), and continuing treatment for depression among VA outpatients (Busch et al., 2004). Recently, the VA published information on seven performance measures, including those developed by the Washington Circle and adapted by the National Committee on Quality Assurance for use in health plans (McKellar & Lie, 2004).

Most states have publicly funded mental health, drug and alcohol treatment systems administered through their Departments of Public Health, Mental Health or Substance Abuse. Across the 50 states, the availability of data to support performance measurement varies widely (McCarty et al., 1998). Some states collect basic information on clients in treatment, while other states also collect detailed encounter data and may, in addition, link their data with information from other state agencies on arrests, employment, or other health care utilization. Such linkages present complex challenges relating to database design, record formats, data reliability, and linkage of data sources, however (Saunders & Heflinger, 2004).

Publicly funded treatment systems often are complex. In Tennessee, for example, substance abuse treatment for adolescents involves three public programs: The Medicaid managed care program, a Medicaid set-aside program reserved for and administered by the child welfare agency for youth in state custody, and the Substance Abuse Prevention and Treatment Block grant from the federal government (Saunders & Heflinger, 2003). Often support for data integration is needed. Delaware, Oklahoma, and Washington all participated in the Integrated Data Base (IDB) project, funded by the Substance Abuse and Mental Health Services Administration (SAMHSA), that integrated Medicaid data and State mental health and substance abuse agency data in their respective states (Coffey et al., 2001). These states have used their data to conduct extensive analyses related to quality of care (Luchansky et al., 2000a; Luchansky et al., 2000b; Wickizer et al., 1994; Wickizer et al., 2000).

States sometimes harness the power of their data for quality improvement initiatives. For example, Oklahoma has developed a Regional Performance Management Report to monitor and improve performance. The aim is achieved by quarterly production of reports that give feedback to providers, consumers, Department administrators, and other stakeholders on key indicators of performance for substance abuse and mental health treatment providers funded in whole or in part by the Oklahoma Department of Mental Health and Substance Abuse Services (Oklahoma Department of Mental Health and Substance Abuse Services, 2003).

Medical records. Health care professionals use these records while providing patient care services to review patient data, or to document their own observations, actions or instructions. Either in traditional paper or computerized forms, these records provide detailed clinical information about the care received by a single patient from a single provider or group of providers. Compared with insurance claims, medical records add a close-up view of clinicians' observations, impressions

and thought processes; patients' tests and results; referrals to other providers; patients' medications; and the clinicians' perspective on patients' reactions to their advice. Although the records give a detailed view of clinical care, if not computerized, these data are more expensive to collect and require intrusion into practice settings. Despite extensive efforts to develop standardized vocabulary and an industry standard to guide software developers (Committee of Data Standards for Patient Safety. Board on Health Care Services, 2003), the use of electronic medical records is not widespread among behavioral health treatment facilities. Thus, using medical records for quality measurement or improvement efforts generally still requires complex data abstraction efforts.

For mental health or substance abuse services, the situation is made more complex. First, there is the need to access records from multiple sites in the specialty mental health and substance abuse sector. Second, it often is necessary to access information on services delivered both through the general medical sector and the specialty substance abuse sector. Moreover, clinical records often are obtained from one source of care, so undercounts of the amount of services will occur when patients receive care at multiple sites. Thus, multiple sources may be needed both to identify those with mental health, alcohol or drug problems and to estimate the amount of care they receive because a single source may lead to underestimates of the alcohol, drug and mental health services that are provided.

For assessment of quality for treatment facilities that specialize in mental health and/or substance abuse, under-reporting of mental health, alcohol or drug problems in the medical record is also an issue. Even though everyone treated in such facilities is there for a behavioral health problem, if the focus is on mental health treatment, then substance use disorders may not be coded accurately. Conversely, if the focus of treatment is on substance abuse, mental health problems may not be included on the records. There may be down coding, however, to a behavioral health condition that is viewed as less stigmatizing or coding to maximize reimbursement. Other issues arise with the use of clinical records in these settings, moreover, such as the lack of full reporting of treatment services delivered. Since many treatment services are not reimbursed on a unit basis, some services may simply not be recorded. For example, services such as ancillary vocational counseling may never appear on the record.

Patient/enrollees surveys. Surveys in the area of mental health and substance abuse are aimed at eliciting individuals' experience of behavioral health treatment, including satisfaction with care, self-reported treatment components or events (such as being offered information about medication side effects), and perceived outcomes of care. One influential survey is part of the Mental Health Statistics Improvement Project (MHSIP) Consumer-Oriented Mental Health Report Card, an initiative begun over 25 years ago and supported by the Center for Mental Health Services. The report card features performance measures that reflect consumer concerns in the domains of access, quality/appropriateness, outcomes, and prevention to be used for assessing the effectiveness of mental health services (Ganju et al., 2004; Teague et al., 1997). The MHSIP Report Card has been implemented extensively in public settings, including 45 states (Smith et al., 2004; Teague et al., 1997). A key source of data for the report card is the MHSIP Consumer Survey, which includes items addressing all of these domains. The survey is designed to focus (though not exclusively) on concerns relevant to consumers with severe mental illness and on outcomes.

Another prominent behavioral health consumer survey is the Experience of Care and Health Outcomes (ECHO) Survey (CAHPS Survey Users Network, 2004). The ECHO Survey evolved from the Consumer Assessment of Behavioral Health Survey (CABHS) (Eisen et al., 1999; Eisen et al., 2001; Shaul et al., 2001), which in turn was developed based upon the CAHPS[®] (originally, the Consumer Assessment of Health Plan Study) Survey designed to collect general health plan consumer data (Eisen et al., 2001). The ECHO development process reflected an effort to develop a standardized instrument that could be used by organizations such as the National Committee for Quality Assurance (NCQA) to develop ratings of behavioral health care and plans for consumers (CAHPS Survey Users Network, 2004; Daniels et al., in press). The ECHO survey includes questions on a wide range of issues including access to care, experience of provider's communication and listening skills, satisfaction with care, and perceived benefits of treatment. In 2002, NCQA adopted a version of the ECHO survey as a HEDIS measure for commercial managed behavioral healthcare organizations (MBHOs), though not for general health plans (CAHPS Survey Users Network, 2004). While the ECHO survey is not a requirement for MBHO accreditation, MBHOs seeking accreditation may use these survey results to meet some MBHO accreditation standards. NCQA encourages managed behavioral health care organizations (MBHOs) to conduct the survey and use it for this purpose (NCQA, 2003).

In addition, there are many other surveys used by various systems and organizations. To improve comparability across organizations while fostering adaptability to a particular organization's population and needs, the Modular Survey Initiative sponsored by the Substance Abuse and Mental Health Services Administration (SAMHSA) aims to develop and promote a common set of perception of care survey tools that can be widely used across settings and populations (Doucette, 2004; Merrick, 2004; Smith, 2004). Specifically, the instrument is being designed to be applicable across groups (youth and adults), service settings (mental health/substance abuse), and delivery/financing vehicles (i.e., public and private, insured and uninsured). Development is currently underway, conducted through the Forum on Performance Measures in Behavioral Health. A modular approach is being used to develop groups of items that address both concerns common to all responding groups and those limited to specific groups. The items will be derived wherever possible from existing behavioral health consumer surveys such as the MHSIP and ECHO surveys, and will include process, outcome and satisfaction measures from the consumer perspective.

Population Issues

Individuals who need mental health or substance abuse services vary in terms of the severity of their illness and cultural issues. These variations might be addressed in performance measures in several ways.

Risk adjustment. Performance measures are influenced by patients' characteristics including severity of the mental health or substance use disorder, other medical conditions, sex, race, and socioeconomic status. For example, Carlson and colleagues (Carlson et al., 2002) found that older and healthier patients, patients with less education, those whose insurance paid all costs of care, women and frequent users rated their behavioral health care and health plan more highly. Wells and colleagues report less access to care, poor quality of care, and greater unmet need for alcoholism, drug abuse, and mental health treatment for Hispanics and African Americans in comparison to whites (Wells et al., 2001). In addition, some research has shown that people with substance abuse disorders often fail to receive regular health care (Saitz R et al., 1997) and can be less likely to receive preventive services and treatment for chronic medical disorders (Cryer et al., 1999; Desai et al., 2002; Druss et al., 2000; Druss et al., 2002; Petersen et al., 2003). This lower

contact with the medical system also can have implications for the quality of behavioral health services.

The implications of these and other studies showing differences in quality of care depend on the type of measure and purpose of the measure. Regardless of specific treatment interventions, outcomes can be influenced by patient factors (e.g. age, diagnosis, severity, co-morbidities), sociodemographic factors (e.g. marital status, socioeconomic status), or patient and family attitudes. Moreover, health plans with a reputation for specialized services to treat specific problems, such as schizophrenia, may attract patients with more severe profiles. In comparing outcomes across health plans, therefore, it is key to adjust for these factors (Pincus et al., 1996). To appropriately target improvement activities, it is necessary to use process measures for continuous quality improvement efforts and to understand the differences in quality among subgroups. Thus, it may be useful to conduct sub-group analyses. If, however, plans are being held to a minimum standard of care and if improvement over time is the goal, then it may be less appropriate to adjust for case mix.

Serious, yet rare, conditions. It is crucial to include large enough numbers of patients for any unit of analysis (e.g., health plan or practice setting) so that performance rates will be stable. The National Committee on Quality Assurance (NCQA) requires a minimum of 30 for the measures in the Health Employer Data and Information Set (HEDIS). For substance abuse or some specific mental health problems, low prevalence may mean that there are not enough patients to meet this minimum standard. For example, with the average substance abuse identification rate of 0.46 percent among commercially insured adults from 2001 data obtained from Medstat, Inc. (Horgan et al., 2004a), a plan would need to have enrollment of 6,500 person years to expect 30 enrollees with substance abuse diagnoses.

Co-occurring mental health and substance abuse conditions. Often mental health and substance abuse problems are co-occurring. Among commercially insured adults, about half of those with substance abuse diagnoses during a year also had mental health diagnoses during the same year. However, among the much larger group of individuals with mental health diagnoses, only about 3 percent also had substance abuse claims. Certainly evidence shows that treating both the mental health and substance abuse conditions often is necessary. For example, Ouimette and colleagues' (Ouimette et al., 2000) analysis suggests that PTSD-focused treatment services are an essential

treatment component for substance abuse/dependence patients with PTSD. Until now, however, most performance measures are focused either on substance abuse or on mental health and have not yet integrated the two.

Data Quality Issues

Data quality issues unique to behavioral health include coding of diagnoses and procedures on administrative data systems, the levels of detail reported in some data systems, and responses to surveys by individuals with behavioral health problems.

Miscoding of Mental Health and Substance Abuse Diagnoses. In commercial data systems, the predominant diagnosis coding system, the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) (National Center for Health Statistics) and insurers' data formats (generally with places for two or three diagnosis codes) hold out the apparent promise of identifying people with addiction problems related to specific drugs, such as addiction to cocaine, or specific serious mental health problems. There is often little incentive, however, to record multiple diagnosis codes or to make codes accurate or specific to the patient's condition. Indeed, the coded diagnosis may be made very general or some other diagnosis may be substituted to protect patients' privacy. Miscoding of behavioral health claims also can include up coding to increase severity or specificity although down coding to substitute broader terms or symptoms is more common (Rost et al., 1994b; Wynia et al., 2000). Reasons include insurability, confidentiality and stigma, and issues about how best to describe patients who do not readily fit into the current diagnostic terms (Rushton et al., 2002).

Failure to Code Mental Health and Substance Abuse Diagnoses. Miscoding of substance abuse or mental health services in claims data also can occur if providers fail to code mental health or substance abuse diagnoses. Even if noted on the medical record, alcohol or drug problems may not appear on the claim because of perceived irrelevance to the medical diagnosis, fear of stigma or lack of space. Adams and colleagues reported that when matched medical records indicated alcoholism, only 77 percent of Medicare claims had an alcohol diagnosis on the Medicare claims. In this case lack of space was apparently the most important source of the mismatch, because the data file only had room for five diagnoses. (Adams et al., 1993). Primary and secondary inpatient mental health diagnoses were reliable, but claims data failed to capture several types of

outpatient services in a comparison between Medicaid claims data and medical record information (Walkup et al., 2000).

Medical records are often held up as the “gold standard” but the issue of miscoding also arises in medical records from general medical settings where mental health problems, as well as drug/alcohol misuse, abuse or dependence often are under recorded. For medical conditions, records are considered to provide more complete clinical information than other sources of information, particularly insurance claims or managed care encounter data. Numerous studies report, however, that patients who have alcohol or drug problems according to structured instruments, such as the Addiction Severity Index (ASI) or biological tests are not necessarily given a substance abuse related diagnosis on their ambulatory (Coulehan et al., 1987) (Moore & Malitz, 1986), clinic (Kirkpatrick et al., 1988), or hospital-based medical charts (Baird et al., 1989). Although the diagnosis may not be provided, sometimes notes may indicate a clinician’s concern about “problem drinking” or excessive use of some other substance. As computerized medical records become more common, it may be possible to do word searches, thus assisting in identifying medical records where substance abuse is recorded in clinical notes but not in diagnostic coding.

Coding of Substance Abuse Procedures. The availability of procedure codes that capture the treatment of behavioral health is particularly problematic for substance abuse because codes for procedures specific to the treatment of addictions are not included in the most widely used coding system for procedures, the Current Procedures Terminology (CPT), published by the American Medical Association and updated annually. Generic psychiatric procedure codes that do not differentiate between mental health and substance abuse services generally are used instead.

The Health Insurance Portability and Accountability Act (HIPAA) contains electronic transaction regulations, which require a national procedure code set. This has the potential to improve this situation of lack of substance abuse treatment codes because not only has HIPAA standardized procedures and durations of treatment, but the substance abuse code set has been expanded to more precisely reflect the wide range of existing types of substance abuse services (SAMHSA, 2004a). Coupled with the addition of new place of service codes, the methodological difficulties in accurately identifying treatment services from administrative patient records should improve.

Not enough detail collected in some settings. Performance measures based on claims or encounter type data generally require detailed information on date of service, types of services, and diagnosis for specific individuals. While some states are leaders in data systems as discussed earlier, many report aggregate utilization statistics or report the start and end dates of episodes of care without specific detail on dates for all services. However, it is encouraging that more states are moving toward the level of detail required for many performance measures and ongoing research efforts have the goal of exploring state capacity to do substance abuse performance measures.

Responses to consumer surveys by persons with behavioral health problems. In designing consumer surveys that focus on behavioral health, it is key to keep in mind some methodological issues that may need special attention:

- **Referent.** It is key to make efforts to distinguish which providers and payers the respondent is considering when answering survey questions. Many respondents may have received behavioral health services from multiple providers (primary care providers, specialty providers, community clinics) and paid for by multiple sources (general managed care plan, behavioral health carve out, block grant funds).
- **Sampling frame.** For managed care plans or managed behavioral health care organizations with defined populations of enrollees, it is possible to use that frame to select a sample to be surveyed. For some purposes (such as asking about screening) it may be useful to sample from the entire population, while for other purposes (such as asking about treatment) it may be useful to sample from the population with behavioral health diagnoses. In public programs, the sample frame may more often be limited to those who received treatment.
- **Timing.** Often, questions relate to the services that respondents have received, such as follow-up after an acute phase of treatment. Thus, surveys might need to be administered on a rolling basis some period of time after episode of care. This timing issue is particularly relevant for surveys that address outcomes since responses may vary depending on whether the respondent has recently begun treatment, is actively engaged in treatment, or is receiving continuing care. In a related concept, recently, McLellan and colleagues (McLellan et al., forthcoming) proposed the concept of “concurrent recovery

monitoring” for outpatient addiction treatment arguing that clinical personnel should collect outcome measures regularly during the course of treatment.

- **Patient factors.** Sometimes, patient factors may influence survey responses without necessarily relating to actual differences in care quality (Hermann et al., 1998). Greenley et al. (Greenley et al., 1982) found that psychologically distressed patients reported greater dissatisfaction with the quality of their health care services. In a study investigating the relationship between psychiatric disorders and satisfaction scores among Medicare beneficiaries, Hermann and colleagues (Hermann et al., 1998) found that a significantly lower proportion of patients with psychiatric disorders reported satisfaction with overall quality of care than patients without psychiatric disorders. After disaggregating components of care, they found that among elderly patients, those with psychiatric disorders were significantly less likely to report satisfaction with their follow-up care and with their doctors’ concern with their overall health.

Calculation Issues

Two issues that are key to calculating performance measures are obtaining the numbers of patients available for statistically significant and meaningful analyses and the definition of new episodes.

Small numbers. In considering performance measures that require information on patients in enrolled populations, such as those enrolled in health plans or Medicaid plans, the issue of small numbers of patients can arise from several sources. First, the prevalence of the condition in the population may be low, as is the case with substance abuse or with specific serious mental illnesses. Second, even if the condition is prevalent, as discussed earlier, failures to code accurately may make the identification of patients through claims data difficult. Third, if the health plan or organization is small, there may be inadequate number of patients for analysis. While NCQA has set 30 as the minimum number for which a measure will be reported, certainly for both statistical significance and meaningfulness, much larger numbers are desirable.

The fragmented service delivery system as well as issues of small numbers of patients with specific mental health or substance use disorders makes the combination of data from multiple sources essential. Moreover, in the very activity of planning such joint efforts, planning the development and use of performance measures can occur. One example is the Massachusetts

Health Quality Partners (MHQP), a broad-based coalition of physicians, hospitals, health plans, purchasers and government agencies established in 1995 to work together to promote improvement in the quality of health care services in Massachusetts. Currently, MHQP is planning to combine HEDIS performance measures, including the measure focused on antidepressant management, from multiple plans which will offer a large enough patient base to allow an examination of quality at the provider group level (MHQP, 2004). In the area of behavioral health, MHQP is collaborating with the state to develop performance incentives for the behavioral health carve-out.

Episodes. Defining a service that begins a new episode of care is essential for many performance measures (Hornbrook et al., 1985) such as the substance abuse initiation and engagement measures or the mental health depression medication measure. The date-specific diagnostic and procedure detail available in claims data make it a rich and efficient source for the development of episodes of care. The nature of this data source allows the sequential treatment for a given condition over time to be linked together to form a single unit of analysis in more cost-efficient manner than what is usually possible with other methods, such as medical record reviews (Wingert et al., 1995).

Regardless of the condition, certain elements are common to all episodes, such as defining the beginning and end point for the episode. A key issue relating to defining the beginning of an episode is being able to distinguish between an incident event and follow-up care, as distinct from a new episode of care. For example, the initiation of substance abuse treatment measure adopted by NCQA requires a 60-day “negative diagnosis period” for new episodes of care. This period without any claims is also called a “clean periods” in the literature on episodes of care. The manner in which new episodes are defined can sometimes create data restrictions. This is especially relevant if the data being used does not support lengthy base periods or the “negative diagnosis period” requirement leads to the loss of potential episodes due to overly restrictive study time windows. In a study of anti-depression medication using claims and pharmacy data, Kerr and colleagues (Kerr et al., 2000) compared using a four month and a nine month “negative diagnosis period” to define a new episode. The different approaches resulted in different numbers of patients eligible to study, but the conclusions about differences between the two groups in depression prescribing were not clear-cut. Similarly, an analysis of adults who were

commercially insured showed that extending the “negative diagnosis period” from 60 to 90 days reduced the number of new substance abuse episodes, as expected, but that the initiation and engagement rates did not change markedly (Horgan et al., 2004a).

ADOPTION AND IMPLEMENTATION OF PERFORMANCE MEASURES

Having well developed and meaningful performance measures is key to driving quality improvement in behavioral healthcare; however, it is not a sufficient condition. As noted previously, significant progress has been made in measure development in recent years. Major developers, such as the National Committee on Quality Assurance’s Health Employer Data and Information Set (HEDIS) and state mental health and substance abuse agencies, have adopted performance measures in both mental health and substance abuse treatment; however, it is how the various stakeholders in behavioral health use these types of measures that ultimately drives quality improvement.

Evidence to date shows ample room for improvement with relatively low rates and scores on many measures. In fact, recent reporting on HEDIS behavioral health measures shows that the measures have been relatively flat over the last few years, unlike many of the medical measures that have shown significant improvement (NCQA, 2004b). Between 1999 and 2002, the five behavioral health measures in HEDIS only had a 2 percentage point increase overall (48 percent to 50 percent) and the two lowest measures actually declined 0.5 percentage points. Whereas during the same time period, the non-behavioral health measures in HEDIS increased 10 percentage points overall (57 percent to 67 percent). The five lowest measures increased 18 percentage points and the five highest measures were up 6 percentage points indicating substantial improvement for the non-behavioral health measures (Goplerud, 2004).

We are at a critical juncture for behavioral health where a foundation has been developed in performance measurement, but stakeholders, including purchasers, health plans, provider groups, individual clinicians, patients, and researchers need to engage in the process of effectively using the measures to drive quality improvement (Horgan et al., 2003; Pincus et al., 2004). Adoption of performance measures involves multiple stakeholders with cross-cutting issues for all. In making this commitment to quality improvement in behavioral health, it is important to be cognizant of these common barriers, as we examine what each stakeholder can contribute in terms

of use of performance measures. For example, multiple measures of the same concept by different developers may exist. Indeed, the plethora of similar measures and the lack of standardization is a major challenge for both health and behavioral health performance measures. Coordination of the considerable activity in this area, an essential next step, is the explicit goal of the Forum on Performance Measures in Behavioral Health and Related Service Systems (Bartlett 2004). This goal is consistent with the aims of the National Quality Forum, which is focused on accomplishing these objectives more broadly throughout the entire health care system (National Quality Forum).

There also may be confusion about the relationship between process and outcome measures and when each is appropriate and sufficient. Recognition that patients with mental and addictive disorders may be difficult and expensive to treat and that they often have needs outside the health care system presents additional challenges. Solutions require collaboration across multiple stakeholders.

Purchasers

Purchasers may lack confidence in the business case for paying for services for mental and addictive disorders. However, they can use performance measures to ensure that what they are paying for is of high quality. Both public and private purchasers have many levers at their disposal to drive quality improvement. They can use performance measures in purchasing decisions, seeking health plans and MBHOs that have better scores on measures. They can select a benefit design that reflects a chronic care approach and incorporates performance measures that cut across the full continuum of care, including maintenance. They can use performance measures in contracts and work with plans to improve scores over time. In particular, performance measures can be linked with financial and non-financial incentives and penalties. Purchasers' "pay for performance" initiatives, which directly reward providers, are beginning to be used in the broader medical area, and have potential for the behavioral health area (Bachman, 2004).

Health Plans

Health plans have legitimate qualms about employing performance measures. The model type (e.g. group/staff HMO or network plan) implies different levels of plan control, and impacts the

plan's ability to influence provider behavior. There are also concerns about both the expense of collecting the performance measures and the expense associated with implementing quality improvement initiatives. In a highly competitive health care marketplace, demand for performance measures from other stakeholders is important to justify the expense.

Health plans and MBHOs have several options for using performance measures. They can use and improve their data monitoring capacity to provide feedback to individual clinicians on their comparative performance on measures. This might be linked with provider incentive programs. They can design programs for quality improvement that could have a direct link to improving measures, e.g. screening protocols that might lead to higher rates of identification of mental and addictive disorders. MBHOs and health plans can improve their capacity to link across data systems.

Clinicians/Provider Group

Major barriers to performance measurement at the practice level relates to inadequate technology for data collection. This varies tremendously depending on setting with providers in some settings, lacking computer capacity. For example, as recently as 2003, 10.5 percent of specialty substance abuse treatment facilities did not have internet access (SAMHSA, 2004c). In addition, small numbers of patients with a particular diagnosis in a provider group setting may make it difficult to calculate stable performance measures.

The practice setting does offer opportunities to creatively engage in activities that can contribute to improvement in performance measures. Clinicians can follow clinical practice guidelines and otherwise hone their skills in treating patients with mental and addictive disorders. Recognition programs can be developed that use performance measure to honor clinicians who are delivering services of particularly high quality. Individual clinicians can participate in quality improvement initiatives in the practice setting.

Consumers/Patient

With the growing acknowledgement of the importance of patient-centered care (New Freedom Commission on Mental Health, 2003), the role of the consumer is important in using performance measures to drive quality. The mental health area has a long and well-established consumer movement; however, there is a much less mature constituency of consumers and family members

to advocate publicly for improved access and outcomes for persons with substance use disorders. There is also little evidence to date, that consumers use published performance measures in their decision-making. However, consumers can use consumer-oriented tools based on performance measures, e.g. NCQA's Quality Compass, to be better become better informed about their health plan. They can participate in community coalitions and collaborate with patient advocacy groups to encourage the use of performance measures for accountability purposes.

Researchers and Developers

The development of performance measures in behavioral health has made great strides in recent years; however, improvement in measures is a continuous process and relies on both the ongoing development of evidence about treatment effectiveness and monitoring of improvements in data quality. Researchers can play key roles in three areas. First, researchers should continue to develop and test new measures, particularly as they relate the incorporation of the chronic care model into treatment practices. Secondly, researchers also can contribute to a better understanding of the links among performance measures, quality of care, and improved clinical and functional outcomes, the ultimate goal of treatment. Third, because the implementation of performance measures is challenging, researchers can contribute by studying the organizational and system factors that result in the effective use of performance measures.

POTENTIAL CHANGES LEADING TO QUALITY IMPROVMENT

We describe two major areas where focused efforts related to performance measurement have the potential to lead to quality improvements in the delivery of services for mental and addictive disorders. The first area relates to using information technology. The second relates to linking performance measures with incentives to mediate provider and consumer behavior. Indeed calls for better use of information technology and aligning payment policies with quality improvement were key recommendations of the previous IOM study, Crossing the Quality Chasm (Institute of Medicine, 2001), which was focused more broadly on the health care system. Here we look at why using these areas to drive quality is important for the treatment of mental and addictive disorders, while acknowledging that the delivery system for these disorders may be even less well positioned to implement them.

Information to Improve Quality

The potential benefits of greater use of information technology to both improve quality of health care and reduce inefficiencies in the system are well described (Hunt et al., 1998; Steinberg, 2003). Information technology holds promise for improvements in the quality of behavioral health services as well. Sophisticated computer-based information technology is available to serve multiple functions for providers: 1) to identify patients at risk of mental health and substance abuse disorders through the use of screening scales, 2) to provide treatment protocols, algorithms and guidelines to support their decision making, 3) to assist with monitoring patients' illnesses, and 4) to generate reminders regarding follow-up visits (Katon et al., 1997). For example, the VA's Mental Health QUERI program has created a data-driven national program to improve quality of care for Veteran's with mental disorders. It has developed a set of computerized clinical reminders to facilitate physician decision-making processes regarding depression at the point of care delivery, including prompts to screen for depression and evaluate positive screens, and reminders suggesting follow-up treatment or referral (Cannon & Allen, 2000; Kramer et al., 2003).

Automated clinical databases and electronic medical records hold promise for more easily monitoring health care quality. Automated databases allow the generation of reports at many different levels within an organization, assist in the timely distribution of these performance and profile reports, and allow the production of clinically focused quality measures (Kerr et al., 2001). The electronic medical record (EMR) system is a key component of organizing the delivery system to maximize the quality of care processes. Rollman et al. (Rollman et al., 2001) found that primary care physicians who received an electronic message regarding their patient's diagnosis of depression after administration of a screening tool and agreed with the diagnosis (65 percent) were more likely to either begin antidepressant therapy or refer to a specialist. A study by Branger et al. (Branger et al., 1992) found that electronic communication facilitates coordination to a greater extent between primary and specialty departments, which may have particular relevance to behavioral health care, given the dominance of managed behavioral health care carve-out arrangements. Supportive information systems can also provide needed education for patients and their families, which may enhance their collaborative involvement. Research with medical disorders demonstrates that patients who more actively manage their illnesses tend to have better outcomes (Wagner et al., 1996). Providers may also print out educational

materials for their patients from the EMR system to provide customized information (Marshall & Chin, 1998).

Despite the benefits of using tools such as electronic medical records, clinical decision support systems and computerized drug entry, their diffusion into medical practice has been slow. A 2002 survey of medical groups with 20 or more physicians found that most practices did not have electronic data systems with basic clinical information (Casalino et al., 2003). These findings were confirmed in a 2003 national survey of physicians by the Commonwealth Foundation which found the highest use of technology (either routinely or occasionally) was for billing (79 percent) and electronic access to test results (59 percent). However, use of other technologies was less common: EMRs and electronic ordering of tests, procedures and drugs (27 percent) and automated patient reminder system (21 percent). One-fourth of physicians said they used electronic clinical support systems, but only 6 percent used them routinely. There was large variation in use of all technologies by practice size. For example, only 13 percent of physicians in solo practice used EMRs. This increased to 23 percent for practices of 2- 9 physicians, 35 percent for practices of 10 – 49 physicians, and 57 percent for physicians in practices of 50 or more (Audet et al., 2004). This study was a general survey of physicians involved in direct care of adults and while specialty was not reported, the study has implications for the delivery of mental health and substance abuse services in general practice. To the extent that mental health practitioners are based in smaller size practices, one might expect that the diffusion of technology is similarly slow.

Many specialty behavioral health treatment organizations are also very early in the adoption cycle for information technology. For example, a national survey of substance abuse treatment facilities conducted prior to February 2003 found that 20 percent did not have any information systems, e-mail or voice mail for their phone systems. However, 30 percent of programs, primarily those that were part of a larger hospital or health system had well-developed information systems. The remaining 50 percent had computerized administrative systems for billing or record keeping that were typically used only by the administrative staff. Less than 2 percent of the facilities contacted had a clinical information system used for conducting or monitoring care, leading to the conclusion that the specialty addiction treatment system was not well positioned to adopt many evidence-based practices (McLellan et al., 2003; McLellan & Meyers, 2004).

It is of interest to ask why the US health care system is not doing better in the use of information technology in light of its much higher use in many other countries. For example, a 2002 survey found that only 17 percent of US primary care practices use electronic medical records, compared countries such as Sweden (90 percent), Netherlands (88 percent), Britain (58 percent), Germany (48 percent), Italy (37 percent), and Ireland (28 percent) (Chin, 2002). A 2003 physician survey in the US (Bachman, 2004) found that the top three barriers to adoption were: costs of both start-up and maintenance, lack of standards, and lack of time to acquire and implement a new system. Barriers to adoption appeared to be of greater concern in smaller practices, leading the authors to conclude that there appears to be a technological divide with respect to information technology driven by practice size (Audet et al., 2004). While the healthcare system more broadly lacks clinical computerized information systems, it appears that the situation, at least for substance abuse, is graver. Major investment in further developing an information technology infrastructure appears essential, if performance measurement is to more fully reach its potential in driving quality.

Incentives to improve quality

Having well-conceived performance measures and the appropriate technological infrastructure to collect and report on them is an important part of accomplishing the goal of improved quality in behavioral healthcare quality. However, as already discussed, it is how the information is used that becomes the other essential component. The data and the technology need to be complemented by stakeholders changing their behavior in ways that drive quality. Recently much has been written about the use of incentives to mediate behaviors so that performance is improved, thus resulting in higher quality and better patient outcomes. (Dudley et al., 2004; Rosenthal et al., 2004; Steinberg, 2003). Some discussion has been specifically focused on the behavioral health area (Bachman, 2004; Join Together, 2003). Many of these incentives focus on financial strategies. The most common are described as “pay for performance” initiatives, but patient steering is also a type of financial incentive. Non-financial incentives include strategies, such as reputational or recognition programs and reduction in administrative burden related to performance. Both types of incentives can be designed to change behaviors at various levels, including the plan, provider group, individual practitioner, or consumer.

Several disciplines, including economics, psychology, and organizational behavior, provide the theoretical underpinnings for the design of incentive arrangements. Much of the literature has focused on the nature of the incentive itself, including to whom it is aimed, domain of performance, financial magnitude, reinforcement approach, perceived attainability etc. (Dudley et al., 2004). Despite the relatively little attention that has been paid to contextual factors, Dudley and colleagues emphasize their likely importance and further describe them as predisposing or enabling. Predisposing factors include variables such as individual provider characteristics or organizational mission. Enabling factors may operate at the organizational level where the structure of care is decided or at the patient level where the individual may facilitate provider actions (Dudley et al., 2004). Consideration of contextual factors is likely critically important for behavioral health because the system is viewed by many as being resource poor. For example, a national survey concluded that the organizational and administrative infrastructures of many specialty substance abuse treatment facilities were inadequate and unstable (McLellan et al., 2003), which is reflected by the 16 percent annual closure rate reported in SAMHSA's census of programs (SAMHSA, 2001). Perhaps, contextual factors might be expected to be more important in a system that is heavily reliant on public sector funding, and that many regard as under financed.

Paying providers for performance relative to defined quality goals is a concept that has received a lot of discussion and some health plans and purchasers are beginning to experiment with these quality-based purchasing approaches. Rosenthal and colleagues described some of these programs and found that they focused on process, structural and patient satisfaction measures, were frequently linked to HEDIS, rewarded performance relative to other providers, and did not focus on internal quality improvement (Rosenthal et al., 2004). Another review of pay for performance approaches in 12 communities found that plans are using a variety of measures with little standardization; providers risk losing extra payments, not their base; and the payment is usually quite modest (Strunk & Hurley, 2004). Experimental evaluations of these incentive programs are lacking. Dudley and colleagues did a comprehensive literature search, identifying only nine randomized controlled trials, eight using performance-based payment and one using reputational incentives. They concluded that there is some evidence that payment and reputational incentives can work but there is little unequivocal data on which to base a design strategy (Dudley et al., 2004). In late 2004, the Joint Commission on Accreditation of Healthcare

Organizations announced a set of principles to guide the development and refinement of pay-for-performance systems (www.jcaho.org).

Ideas related to using economic incentives to promote quality-improvement initiatives along other dimensions have been proposed. These include: tying provider incentive payments to investment in information technology designed to promote quality improvement, using variable co-payments to encourage consumers to choose providers with better quality profiles, creating tax incentives for providers to invest in quality related information technology, and reducing malpractice premiums or lowering claim ceilings for providers that use specified quality-improvement infrastructures (Steinberg, 2003).

Most pay for performance programs are aimed at general health care; however, a few programs targeted at behavioral health, most often for depression care, are under development (Bachman, 2004). There is also potential for using these types of programs in the public sector. For example, Delaware's Division of Substance Abuse and Mental Health through its contracts with providers has begun a payment system that rewards three components of outpatient substance abuse treatment: engagement/utilization, active participation and program completion (Shapiro, 2004).

Although pay for performance programs in behavioral health are nascent, there are several aspects of behavioral health care delivery that underscore the potential for quality measurement in behavioral health because some of the ingredients on which it could be built already exist. For example, in the private sector there has been strong interest in evaluating provider performance. A 1999 national survey of managed care organizations about the provision of mental health and substance abuse services found that 70 percent of plans used patient satisfaction surveys, 73 percent used performance indicators, and 49 percent evaluated clinical outcomes (Merrick et al., 2002). Specification of performance guarantees are common in purchaser contracts with MBHOs, although typically MBHOs are at risk for penalties if standards are not met, rather than additional payments if performance is greater than expectations (Garnick et al., 2001). Another key aspect of pay for performance is the ability to link with payment mechanisms. It appears that claims and billing systems are in existence in most behavioral health provider settings, even if other technological infrastructure is missing, thus part of what is needed to implement pay for performance programs is in place.

It is clear that well designed evaluations of pay for performance programs are needed (Dudley et al., 2004; Rosenthal et al., 2004). A number of concerns requiring further study are: the need to have a better sense of the size and scope of the incentive necessary to motivate change, whether the raise in quality will come from providers who are already doing well, and the potential or selection against treating the more severely ill or less adherent patients (Rosenthal et al., 2004). Additional concerns have been raised about provider resistance to such strategies (Hermann et al., 2000). Although the jury is still out on the utility of incentives for driving quality in behavioral health, there is evidence that with proper financial and technical investment, these programs could contribute to making progress in behavioral health quality improvement.

CONCLUSION

Performance measurement is a tool that can be used in many ways and at many levels to improve the treatment of mental and addictive disorders. Performance measures can be used in a process of establishing baselines within and across entities at different levels within the health care system, leading to continuous problem solving and quality improvement. The measures themselves must undergo a constant process of evaluation and improvement, as must the supporting technologic infrastructure. Stakeholders can use performance measures in various ways to drive quality, ranging from public reporting mechanisms, provider feedback, and financial incentives, to patient and provider education. The dual processes of performance measure improvement and creative use of performance measures by multiple stakeholders can aid in crossing the quality chasm for behavioral health. It is important to note, however, that the chasm may be larger and more challenging to bridge in behavioral health because of inherent differences in how mental and addictive services are financed, organized and managed which have generally resulted in an underdeveloped information infrastructure.

REFERENCES

- AcademyHealth. (2004). *Health Outcomes Core Library Project* (No. Order No. P.O. 467-MZ-301222). Washington, DC: National Information Center on Health Services Research and Health Care Technology, National Library of Medicine.
- Adams, W. L., Yuan, Z., Barboriak, J. J., & Rimm, A. A. (1993). Alcohol-related hospitalizations of elderly people. Prevalence and geographic variation in the United States. *Jama*, *270*(10), 1222-1225.
- Agency for Healthcare Research and Quality. (2004). *National Quality Measures Clearinghouse*. September 20, 2004, from http://www.qualitymeasures.ahrq.gov/resources/measure_use.aspx#attributes
- AHCMA. (2003). *Turning Knowledge Into Practice: A Manual for Behavioral Health Administrators About Understanding and Implementing Evidence-Based Practices*: American College of Mental Health Administration www.ACMHA.ORG
- American Psychiatric Association. (1993). Practice guideline for major depressive disorder in adults. *Am J Psychiatry*, *150*(4 Suppl), 1-26.
- Anthony, W. (1993). Recovery From Mental Illness: The Guiding Vision of the Mental Health Service System in the 1990s. *Psychological Rehabilitation Journal*, *16*, 11-24.
- Audet, A. M., Doty, M., Peugh, J., Shamasdin, J., Zapert, K., & Schoenbaum, S. (2004). Information Technologies: When Will They Make It into Physicians' Black Bags? *Medscape General Medicine*, December.
- Babor, T. F., & Higgins-Biddle, J. C. (2003). *Quality Measures for Alcohol Misuse and Dependence*: Prepared for the Foundation for Accountability.
- Bachman, J. (2004). "Pay for Performance" in the Behavioral Healthcare Arena. December 19, 2004, from http://www.wpic.pitt.edu/dppc/journalwatch_2004_11.htm
- Baird, M., Burge, S., & Grant, W. (1989). A scheme for determining the prevalence of alcoholism in hospitalized patients. *Alcohol Clin Exp Res*, *13*(6), 782-785.
- Balas, E. A., & Boren, S. A. (2000). *Managing Clinical Knowledge for Health Care Improvement. Yearbook of Medical Informatics*. Bethesda, MD: National Library of Medicine.

- Bartlett, J. (2004, April 19-20). *The Forum on Performance Measures: Finding Commonality through Building Consensus*. Paper presented at the 2nd Annual Forum on Performance Measurement in Behavioral Health and Related Service Systems, Atlanta, GA.
- Bijl, R. V., de Graaf, R., Hiripi, E., Kessler, R. C., Kohn, R., Offord, D. R., et al. (2003). The prevalence of treated and untreated mental disorders in five countries. *Health Aff (Millwood)*, 22(3), 122-133.
- Bodenheimer, T., Wagner, E. H., & Grumbach, K. (2002). Improving primary care for patients with chronic illness. *Jama*, 288(14), 1775-1779.
- Branger, P. J., van der Wouden, J. C., Schudel, B. R., Verboog, E., Duisterhout, J. S., van der Lei, J., et al. (1992). Electronic communication between providers of primary and secondary care. *Bmj*, 305(6861), 1068-1070.
- Busch, S. H., Leslie, D., & Rosenheck, R. (2004). Measuring quality of pharmacotherapy for depression in a national health care system. *Med Care*, 42(6), 532-542.
- CAHPS Survey Users Network. (2004). *Development of the ECHO® Survey by Harvard Medical School*. January 8, 2005, from <http://www.cahpsun.org/Products/ECHO/ECHODEvel.asp>
- Cannon, D. S., & Allen, S. N. (2000). A comparison of the effects of computer and manual reminders on compliance with a mental health clinical practice guideline. *J Am Med Inform Assoc*, 7(2), 196-203.
- Carlson, M. J., Shaul, J. A., Eisen, S. V., & Cleary, P. D. (2002). The influence of patient characteristics on ratings of managed behavioral health care. *J Behav Health Serv Res*, 29(4), 481-489.
- Casalino, L., Gillies, R. R., Shortell, S. M., Schmittdiel, J. A., Bodenheimer, T., Robinson, J. C., et al. (2003). External incentives, information technology, and organized processes to improve health care quality for patients with chronic diseases. *Jama*, 289(4), 434-441.
- Charbonneau, A., Rosen, A. K., Owen, R. R., Spiro, A., 3rd, Ash, A. S., Miller, D. R., et al. (2004). Monitoring depression care: in search of an accurate quality indicator. *Med Care*, 42(6), 522-531.
- Chin, T. (2002). *Americans trail much of Europe in adopting electronic medical records*. from <http://www.ama-assn.org/amednews/2002/09/02/bisf0902.htm>
- Chinman, M., Young, A. S., Rowe, M., Forquer, S., Knight, E., & Miller, A. (2003). An instrument to assess competencies of providers treating severe mental illness. *Ment Health Serv Res*, 5(2), 97-108.

Cochrane Depression Anxiety and Neurosis Group. (2004). *Abstracts of Cochrane Reviews*. January 4, 2004, from <http://www.cochrane.org/cochrane/revabstr/DEPRESSNAbstractIndex.htm>

Cochrane Drugs and Alcohol Group. (2004). *Abstracts of Cochrane Reviews*. December 14, 2004, from <http://www.cochrane.org/cochrane/revabstr/ADDICTNAbstractIndex.htm>

Coffey, R., Grave, r. L., & Shroeder, D. (2001). *Mental Health and Substance Abuse Treatment: Results from a Study Integrating Data from State Mental Health, Substance Abuse, and Medicaid Agencies*. (No. SAMHSA Publication No. SMA-01-3528). Rockville, MD: Center for Substance Abuse Treatment and Center for Mental Health Services, Substance Abuse and Mental Health Services Administration.

Committee of Data Standards for Patient Safety. Board on Health Care Services. (2003). *Key Capabilities of an Electronic Health Record System*. Institute of Medicine of the National Academy of Sciences. www.nap.edu

Coulehan, J. L., Zettler-Segal, M., Block, M., McClelland, M., & Schulberg, H. C. (1987). Recognition of alcoholism and substance abuse in primary care patients. *Arch Intern Med*, 147(2), 349-352.

Cryer, P. C., Jenkins, L. M., Cook, A. C., Ditchburn, J. S., Harris, C. K., Davis, A. R., et al. (1999). The use of acute and preventative medical services by a general population: relationship to alcohol consumption. *Addiction*, 94(10), 1523-1532.

Curie, C., Administrator, Substance Abuse and Mental Health Services Administration, & U.S. Department of Health and Human Services. (2004). Testimony before the House Appropriations Subcommittee on Labor, Health and Human Services, and Education.

Daniels, A., Shaul, J., Greenburg, P., & Cleary, P. (in press). The Experience of Care and Health Outcomes Survey (ECHOTM): A consumer survey to collect ratings of treatment, outcomes and plans. In M. Maruish (Ed.), *The Use of Psychological Testing For Treatment Planning and Outcomes Assessment (3rd ed.)*. Vol. III. *Adult Assessment Instrumentation*. Mahwah, NJ: Lawrence Erlbaum Associates.

Desai, M. M., Rosenheck, R. A., Druss, B. G., & Perlin, J. B. (2002). Mental disorders and quality of diabetes care in the veterans health administration. *Am J Psychiatry*, 159(9), 1584-1590.

Donabedian, A. (1980). *The Definition of Quality and Approaches to Assessment*. Ann Arbor, MI: Health Administration Press.

- Doucette, A. (2004, April 19-20). *Capturing Commonality: developing the Modular Survey*. Paper presented at the 2nd Annual Forum on Performance Measurement in Behavioral Health and Related Service Systems, Atlanta, GA.
- Drake, R. E., Goldman, H. H., Leff, H. S., Lehman, A. F., Dixon, L., Mueser, K. T., et al. (2001). Implementing evidence-based practices in routine mental health service settings. *Psychiatr Serv*, 52(2), 179-182.
- Druss, B. G., Bradford, D. W., Rosenheck, R. A., Radford, M. J., & Krumholz, H. M. (2000). Mental disorders and use of cardiovascular procedures after myocardial infarction. *Jama*, 283(4), 506-511.
- Druss, B. G., Rosenheck, R. A., Desai, M. M., & Perlin, J. B. (2002). Quality of preventive medical care for patients with mental disorders. *Med Care*, 40(2), 129-136.
- Dudley, R., Frolich, A., Robinowitz, D., Talavera, J. A., Peter Broadhead, P., & Harold S. Luft, H. S. (2004). *Strategies To Support Quality-based Purchasing: A Review of the Evidence. Summary, Technical Review Number 10*. Agency for Healthcare Research and Quality, Rockville, MD., 2004, from <http://www.ahrq.gov/clinic/epcsums/qpurchsum.htm>
- ECHO Development Team. (2002). *Experience of Care and Health Outcomes Survey (ECHO)*. November 9, 2004, from <http://www.hcp.med.harvard.edu/echo/home.html>
- Edlund, M. J., Young, A. S., Kung, F. Y., Sherbourne, C. D., & Wells, K. B. (2003). Does satisfaction reflect the technical quality of mental health care? *Health Serv Res*, 38(2), 631-645.
- Eisen, S. V., Shaul, J. A., Clarridge, B., Nelson, D., Spink, J., & Cleary, P. D. (1999). Development of a consumer survey for behavioral health services. *Psychiatr Serv*, 50(6), 793-798.
- Eisen, S. V., Shaul, J. A., Leff, H. S., Stringfellow, V., Clarridge, B. R., & Cleary, P. D. (2001). Toward a national consumer survey: evaluation of the CABHS and MHSIP instruments. *J Behav Health Serv Res*, 28(3), 347-369.
- Fiorentine, R., & Anglin, M. D. (1996). More is better: counseling participation and the effectiveness of outpatient drug treatment. *J Subst Abuse Treat*, 13(4), 341-348.
- Fleming, M. F., Manwell, L. B., Barry, K. L., Adams, W., & Stauffacher, E. A. (1999). Brief physician advice for alcohol problems in older adults: a randomized community-based trial. *J Fam Pract*, 48(5), 378-384.

- Fortney, J., Rost, K., Zhang, M., & Pyne, J. (2001). The relationship between quality and outcomes in routine depression care. *Psychiatr Serv*, 52(1), 56-62.
- French, M. T., Dunlap, L. J., Roman, P. M., & Steele, P. D. (1997). Factors that influence the use and perceptions of employee assistance programs at six worksites. *J Occup Health Psychol*, 2(4), 312-324.
- Friedmann, P. D., Lemon, S. C., Stein, M. D., Etheridge, R. M., & D'Aunno, T. A. (2001). Linkage to medical services in the Drug Abuse Treatment Outcome Study. *Med Care*, 39(3), 284-295.
- Friedmann, P. D., McCullough, D., Chin, M. H., & Saitz, R. (2000). Screening and intervention for alcohol problems. A national survey of primary care physicians and psychiatrists. *J Gen Intern Med*, 15(2), 84-91.
- Ganju, V., Beall, M. A., Callahan, N., Dailey, W., Dumont, J., Hornik, J., et al. (1996). *The MHSIP Consumer-Oriented Mental Health Report Card*. Rockville, MD: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration.
- Ganju, V., Smith, M. E., Adams, N., Allen, J., Bible, J., Danforth, M., et al. (2004). *The MHSIP Quality Report: Next Generation of Mental Health and Performance Measures*.
- Garnick, D. W., Hodgkin D, & Horgan , C. M. (2002a). Selecting data sources for substance abuse services research. *J Subst Abuse Treat*, 22(1), 11-22.
- Garnick, D. W., Horgan, C. M., Hodgkin, D., Merrick, E. L., Goldin, D., Ritter, G., et al. (2001). Risk transfer and accountability in managed care organizations' carve-out contracts. *Psychiatr Serv*, 52(11), 1502-1509.
- Garnick, D. W., Lee, M. T., Chalk, M., Gastfriend, D., Horgan, C. M., McCorry, F., et al. (2002b). Establishing the Feasibility of Performance Measures for Alcohol and Other Drugs. *Journal of Substance Abuse Treatment*, 23, 375-385.
- Gerstein, D. R., & Harwood, H. J. (Eds.). (1990). *Treating Drug Problems: Volume 1*. Washington, D.C.: Institute of Medicine, National Academies Press.
- Goplerud, E. (2004). *Behavioral Health Quality: Who's Paying Attention? : Ensuring Solutions to Alcohol Problems*, The George Washington University.
- Greenley, J. R., Young, T. B., & Schoenherr, R. A. (1982). Psychological distress and patient satisfaction. *Med Care*, 20(4), 373-385.

- Grella, C. E., Hser, Y. I., Joshi, V., & Anglin, M. D. (1999). Patient histories, retention, and outcome models for younger and older adults in DATOS. *Drug Alcohol Depend*, 57(2), 151-166.
- Hermann, R. C., Ettner, S. L., & Dorwart, R. A. (1998). The influence of psychiatric disorders on patients' ratings of satisfaction with health care. *Med Care*, 36(5), 720-727.
- Hermann, R. C., Finnerty, M., Provost, S., Palmer, R. H., Chan, J., Lagodmos, G., et al. (2002a). Process measures for the assessment and improvement of quality of care for schizophrenia. *Schizophr Bull*, 28(1), 95-104.
- Hermann, R. C., Leff, S. H., & Lagodmos, G. (2002b). *Selecting Proces Measures for Quality Improvement in Mental Healthcare, 2002*. The Evaluation Center@HSRI. Retrieved December 5, 2004, from www.hsri.org
- Hermann, R. C., & Palmer, R. H. (2002). Common ground: a framework for selecting core quality measures for mental health and substance abuse care. *Psychiatr Serv*, 53(3), 281-287.
- Hermann, R. C., Regner, J. L., Erickson, P., & Yang, D. (2000). Developing a quality management system for behavioral health care: the Cambridge Health Alliance Experience. *Harv Rev Psychiatry*, 8(5), 251-260.
- Hodgkin, D., Horgan, C. M., & Garnick, D. W. (1997). Make or buy: HMOs' contracting arrangements for mental health care. *Adm Policy Ment Health*, 24(4), 359-376.
- Hodgkin, D., Horgan, C. M., Garnick, D. W., Merrick, E. L., & Goldin, D. (2000). Why carve out? Determinants of behavioral health contracting choice among large U.S. employers. *J Behav Health Serv Res*, 27(2), 178-193.
- Horgan, C., Garnick, D., Renner, P., & Mardon, R. (2004a, October 7). *Performance Measures for Substance Abuse: First-Year HEDIS Results*. Paper presented at the Addiction Health Services Research Conference, Philadelphia, PA.
- Horgan, C., & Levine, H. (1998). The substance abuse treatment system: What does it look like and whom does it serve? In S. Lamb, Greenlick, M & D. McCarty (Eds.), *Bridging the Gap Between Practice and Research*, . Washington, DC. 1998.: Institute of Medicine, National Academy Press,.
- Horgan, C., Merrick, E., Garnick, D., Hodgkin, D., Cenczyk, R., Lusenhop, R., et al. (2003). *Provision of Mental Health Services in Managed Care Organizations*. Rockville, MD: US. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.

- Horgan , C. M. (1997). Need and Access to Drug Treatment. In *Treating Drug Problems Effectively*. Commissioned by Milbank Memorial Fund and NIDA,; Blackwell Press.
- Horgan , C. M., Garnick, D. W., Merrick, E., Hodgkin, D., & Volpe-Vartanian, J. (2004b). *Substance Abuse and Mental Health Services in Health Plans: Results from a Nationwide Study. Report to Respondents*. Waltham, MA: Heller School for Social Policy and Management, Brandeis University.
- Horgan , C. M., & Hodgkin, D. (1999). Mental Illness and Substance Abuse Among Low Income Populations. In M. Lillie-Blanton (Ed.), *Access to Health Care: Promises and Prospects for Low Income Americans*. Baltimore, MD: Johns Hopkins University Press.
- Horgan , C. M., McCorry, F., & W, G. D. (2003). *Creating Policy Relevant Information: The Case of Washington Circle*. Paper presented at the AcademyHealth Annual Meeting, Nashville TN.
- Horgan , C. M., & Merrick, E. (2001). Financing of Substance Abuse Treatment Services. In G. M (Ed.), *Alcoholism: Volume 15 Services Research in the Era of Managed Care: Organization, Access, Economics Outcome*. New York, NY: Kluwer Academic/Plenus Publishers.
- Horgan , C. M., W, G. D., Merikangas, K. R., D, H., Sciegaj, M., & Goldin, D. (2000). Structuring Behavioral Healthcare: Carved Out or Integrated? *Compensation and Benefits Management*, 16(4), 39-46.
- Hornbrook, M. C., Hurtado, A. V., & Johnson, R. E. (1985). Health care episodes: definition, measurement and use. *Med Care Rev*, 42(2), 163-218.
- Hubbard, R. L., Craddock, S. G., & Anderson, J. (2003). Overview of 5-year follow up outcomes in the drug abuse treatment outcome studies (DATOS). *J Subst Abuse Treat*, 25(3), 125-134.
- Humphreys, K., & Weingardt, K. R. (2000). Assessing readmission to substance abuse treatment as an indicator of outcome and program performance. *Psychiatr Serv*, 51(12), 1568-1569.
- Humphreys, K., & Weisner, C. (2000). Use of exclusion criteria in selecting research subjects and its effect on the generalizability of alcohol treatment outcome studies. *Am J Psychiatry*, 157(4), 588-594.
- Hunt, D. L., Haynes, R. B., Hanna, S. E., & Smith, K. (1998). Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. *Jama*, 280(15), 1339-1346.

- Institute for Clinical Systems Improvement (ICSI). (2004). *Major depression in adults in primary care*. December 16, 2004, from <http://www.icsi.org/knowledge/detail.asp?catID=29&itemID=179>
- Institute of Medicine. (1990a). *Broadening the Base of Treatment for Alcohol Problems*. Washington, D.C.: National Academies Press.
- Institute of Medicine. (1990b). *Treating Drug Problems: Volume 1*. Washington, DC: National Academies Press.
- Institute of Medicine. (1997). *Managing Managed Care: Quality Improvement in Behavioral Health*. Washington Dc: National Academies Press.
- Institute of Medicine. (1998). *Bridging the Gap Between Practice and Research: Forging Partnerships with Community -Based Drug and Alcohol Treatment*. Washington DC: National Academies Press.
- Institute of Medicine. (2001). *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, D.C.: INational Academies Press. <http://books.nap.edu/books/0309072808/html/2.html#pagetop>
- Institute of Medicine (Ed.). (1996). *Primary Care: America's Health in a New Era*. Washington, D.C.: Institute of Medicine, National Academies Press.
- Join Together. (2003). *Rewarding Results - Improving the Quality of Treatment for People with Alcohol and Drug problems*. from www.jointogether.org
- Joint Commission on Accreditation of Healthcare Organizations. *How to Select a Performance Measurement System*. September 20, 2004, from <http://www.jcaho.org/accredited+organizations/behavioral+health+care/oryx/system+selection/>
- Katon, W., Von Korff, M., Lin, E., Unutzer, J., Simon, G., Walker, E., et al. (1997). Population-based care of depression: effective disease management strategies to decrease prevalence. *Gen Hosp Psychiatry*, 19(3), 169-178.
- Kazel, R. (2004). Aetna buys back its mental health division. *American Medical News*. <http://www.ama-assn.org/amednews/2004/12/27/bisd1227.htm>
- Kerr, E. A., Krein, S. L., Vijan, S., Hofer, T. P., & Hayward, R. A. (2001). Avoiding pitfalls in chronic disease quality measurement: a case for the next generation of technical quality measures. *Am J Manag Care*, 7(11), 1033-1043.

- Kerr, E. A., McGlynn, E. A., Van Vorst, K. A., & Wickstrom, S. L. (2000). Measuring antidepressant prescribing practice in a health care system using administrative data: implications for quality measurement and improvement. *Jt Comm J Qual Improv*, 26(4), 203-216.
- Kirkpatrick, B., Johnson, M. S., Earp, J. A., & Fletcher, R. H. (1988). Accuracy of chart diagnoses of alcoholism in patients with a history of psychosis. *Psychiatr Med*, 6(1), 65-71.
- Kramer, T. L., Owen, R. R., Cannon, D., Sloan, K. L., Thrush, C. R., Williams, D. K., et al. (2003). How well do automated performance measures assess guideline implementation for new-onset depression in the Veterans Health Administration? *Jt Comm J Qual Saf*, 29(9), 479-489.
- Lee, M. T., Garnick, D. W., Miller, K., & Horgan, C. M. (2004). Datapoints: adolescents with substance abuse: are health plans missing them? *Psychiatr Serv*, 55(2), 116.
- Lehman, A. F., Kreyenbuhl, J., Buchanan, R. W., Dickerson, F. B., Dixon, L. B., Goldberg, R., et al. (2004). The Schizophrenia Patient Outcomes Research Team (PORT): updated treatment recommendations 2003. *Schizophr Bull*, 30(2), 193-217.
- Lehman, A. F., & Steinwachs, D. M. (1998). Patterns of usual care for schizophrenia: initial results from the Schizophrenia Patient Outcomes Research Team (PORT) Client Survey. *Schizophr Bull*, 24(1), 11-20; discussion 20-32.
- Leslie, D. L., & Rosenheck, R. A. (2001). Use of pharmacy data to assess quality of pharmacotherapy for schizophrenia in a national health care system: individual and facility predictors. *Med Care*, 39(9), 923-933.
- Lima, M. S., Reisser, A. A. P., Soares, B. G. O., & Farrell, M. (2004). *Antidepressants for cocaine dependence (Cochrane Review)*. In: *The Cochrane Library, Issue 4, 2004*. Chichester, UK: . John Wiley & Sons, Ltd.
<http://www.cochrane.org/cochrane/revabstr/AB002950.htm>
- Luchansky, B., Brown, M., Longhi, D., Stark, K., & Krupski, A. (2000a). Chemical dependency treatment and employment outcomes: results from the 'ADATSA' program in Washington State. *Drug Alcohol Depend*, 60(2), 151-159.
- Luchansky, B., He, L., Krupski, A., & Stark, K. D. (2000b). Predicting readmission to substance abuse treatment using state information systems. The impact of client and treatment characteristics. *J Subst Abuse*, 12(3), 255-270.

- Mark, T. L., Coffey, R. M., King, E., Harwood, H., McKusick, D., Genuardi, J., et al. (2000). Spending on mental health and substance abuse treatment, 1987-1997. *Health Aff (Millwood)*, 19(4), 108-120.
- Marshall, P. H., & Chin, H. (1998). *The effects of an electronic medical record on patient care: clinician attitudes in a large HMO*. Paper presented at the Proc AMIA Symp.
- Matevia, M. L., Poon, D., Goldman, W., Cuffel, B., & McCulloch, J. (2001). Access to network clinicians in a managed behavioral health organization. *Psychiatr Serv*, 52(11), 1428.
- McCarty, D., & Goldman, H. H. (2004). Treatment for Alcohol and Drug Dependence: History, Workforce, Organization, Financing, and Emerging Policies. In J. H. Lowinson, P. ERuiz, R. P. Millman & J. G. Langwood (Eds.), *Substance Abuse: A Comprehensive Textbook, 4th Edition* (pp. 1346-1360). Philadelphia, PA: Lippincott, Williams, and Wilkins.
- McCarty, D., McGuire, T. G., Harwood, H. J., & Field, T. (1998). Using state information systems for Drug Abuse Services Research. *American Behavioral Scientist*, 41(9), 1090-1106.
- McCorry, F., Garnick, D. W., Bartlett, J., Cotter, F., & Chalk, M. (2000). Developing performance measures for alcohol and other drug services in managed care plans. *The Joint Commission Journal on Quality Improvement*, 26(11), 633-643.
- McGlynn, E. A. (1998). Choosing and evaluating clinical performance measures. *Jt Comm J Qual Improv*, 24(9), 470-479.
- McGlynn, E. A., Asch, S. M., Adams, J., Keesey, J., Hicks, J., DeCristofaro, A., et al. (2003). The quality of health care delivered to adults in the United States. *N Engl J Med*, 348(26), 2635-2645.
- McKellar, J. D., & Lie, C.-C., Saweikis Meghan. (2004). *VA Care for Substance Use Disorder Patients: Indicators of Facility and VISN Performance (Fiscal Years 2002 and 2003)*. Palo Alto, CA: Program Evaluation and Resource Center and HSR&D Center for Health Care Evaluation www.chce.info
- McLellan, A., McKay JR, Forman R, Cacciola, J., & Kemp, K. (forthcoming). Reconsidering the Evaluation of Addiction Treatment: From Retrospective Follow-Up to Concurrent Recovery Monitoring. *Addiction*.
- McLellan, A. T., Arndt, I. O., Metzger, D. S., Woody, G. E., & O'Brien, C. P. (1993). The effects of psychosocial services in substance abuse treatment. *Jama*, 269(15), 1953-1959.

- McLellan, A. T., Carise, D., & Kleber, H. D. (2003). Can the national addiction treatment infrastructure support the public's demand for quality care? *J Subst Abuse Treat*, 25(2), 117-121.
- McLellan, A. T., Hagan, T. A., Levine, M., Gould, F., Meyers, K., Bencivengo, M., et al. (1998). Supplemental social services improve outcomes in public addiction treatment. *Addiction*, 93(10), 1489-1499.
- McLellan, A. T., & Hunkeler, E. (1998). Patient satisfaction and outcomes in alcohol and drug abuse treatment. *Psychiatr Serv*, 49(5), 573-575.
- McLellan, A. T., Lewis, D. C., O'Brien, C. P., & Kleber, H. D. (2000). Drug dependence, a chronic medical illness: implications for treatment, insurance, and outcomes evaluation. *Jama*, 284(13), 1689-1695.
- McLellan, A. T., & Meyers, K. (2004). Contemporary addiction treatment: a review of systems problems for adults and adolescents. *Biol Psychiatry*, 56(10), 764-770.
- Merrick, E., Garnick, D. W., Horgan, C. M., & Hodgkin, D. (2002). Quality measurement and accountability for substance abuse and mental health services in managed care organizations. *Med Care*, 40(12), 1238-1248.
- Merrick, E., Horgan, C. M., W, G. D., D, H., Keck, K., & Cenczyk, R. (2003). The EAP/Behavioral Health Carve-Out Connection. *Employee Assistance Quarterly*, 18(3), 1-13.
- Merrick, E. L. (2004, April 19-20). *The Modular Survey and Consumer Voices in Substance Abuse Treatment*. Paper presented at the 2nd Annual Forum on Performance Measurement in Behavioral Health and Related Service Systems, Atlanta, GA.
- MHQP. (2004). *Massachusetts Health Quality Partners*. from www.mhqp.org
- MHSIP Task Force. (1996). *The Final Report of the Mental Health Statistics Improvement Program (MHSIP) Task Force on Consumer-Oriented Mental Health Report Card*. Washington, D.C.: Center for Mental Health Services, Substance Abuse and Mental Health Services Administration www.mhsip.org
- MHSIP Task Force. (2004). *Mental Health Statistics Improvement Project (MHSIP)*. December 6, 2004, from <http://www.mhsip.org/>
- Miller, W. R., & Wilbourne, P. L. (2002). Mesa Grande: a methodological analysis of clinical trials of treatments for alcohol use disorders. *Addiction*, 97(3), 265-277.

- Moore, R. D., & Malitz, F. E. (1986). Under diagnosis of alcoholism by residents in an ambulatory medical practice. *J Med Educ*, 61(1), 46-52.
- National Center for Health Statistics. *Classifications of Diseases and Functioning & Disability . International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)*. November 30, 2004, from <http://www.cdc.gov/nchs/about/otheract/icd9/abtcd9.htm>
- National Quality Forum. December 16, 2004, from <http://www.qualityforum.org>
- NCQA. (1998). *National Committee on Quality Assurance: Desirable Attributes of HEDIS Measures*. National Committee on Quality Assurance. Retrieved September 20, 2004, from <http://www.ncqa.org/Programs/HEDIS/desirable%20attributes.html>
- NCQA. (2003). *Survey Provides First Look at Satisfaction With Managed Behavioral Health Care Services*. National Committee on Quality Assurance. Retrieved January 8, 2005, from <http://www.ncqa.org/communications/news/echoresults.htm>
- NCQA. (2004a). *HEDIS -- Health Plan Employer Data and Information Set, Volume 2 Technical Specifications*. Washington, D.C.: National Committee on Quality Assurance. www.ncqa.org
- NCQA. (2004b). *The State of Health Care Quality 2004*. National Committee on Quality Assurance. Retrieved, from www.ncqa.org
- Network for the Improvement of Addiction Treatment (NIATx). Substance Abuse and Mental Health Services Administration and Robert Wood Johnson Foundation. Retrieved November 29, 2004, from www.niatx.org
- New Freedom Commission on Mental Health. (2003). *Achieving the Promise: Transforming Mental Health Care in America, Final Report* (No. DHHS Pub. No. SMA-03-3832). Rockville, MD.
- NIAAA. (2000). *Special Report to the U.S. Congress on Alcohol and Health: Highlights from Current Research* (No. NIH Publication No. 00-1583): National Institute on Alcohol Abuse and Alcoholism.
- NIDA. (2004). Report of the Blue Ribbon Task Force on Health Services Research at the National Institute on Drug Abuse.
- Oklahoma Department of Mental Health and Substance Abuse Services. (2003). *Early Warning System for Performance Monitoring and Improvement May*.
- Organization for Economic Co-operation and Development. (2003). *The OECD Health Care Quality Indicators Project. Report of the Mental Health Panel*. New York.

- Ouimette, P. C., Moos, R. H., & Finney, J. W. (2000). Two-year mental health service use and course of remission in patients with substance use and posttraumatic stress disorders. *J Stud Alcohol*, *61*(2), 247-253.
- Petersen, L. A., Normand, S. L., Druss, B. G., & Rosenheck, R. A. (2003). Process of care and outcome after acute myocardial infarction for patients with mental illness in the VA health care system: are there disparities? *Health Serv Res*, *38*(1 Pt 1), 41-63.
- Physician Consortium for Performance Improvement. *Clinical Performance Measures: Major Depressive Disorder*. American Medical Association (AMA) Division of Clinical Quality Improvement Web site: Retrieved December 11, 2004, from www.ama-assn.org/go/quality
- Pincus, H., Pechora, C., & Defray, F. (2004). *Depression in Primary Care: Linking Clinical and System Strategies*. Paper presented at the Robert Wood Johnson Foundation Depression in Primary Care (DPC) Annual Grantee Meeting, Huntington Beach, CA.
- Pincus, H. A., Zarin, D. A., & West, J. C. (1996). Peering into the 'black box'. Measuring outcomes of managed care. *Arch Gen Psychiatry*, *53*(10), 870-877.
- Regier, D. A., Farmer, M. E., Rae, D. S., Myers, J. K., Kramer, M., Robins, L. N., et al. (1993a). One-month prevalence of mental disorders in the United States and sociodemographic characteristics: the Epidemiologic Catchment Area study. *Acta Psychiatr Scand*, *88*(1), 35-47.
- Regier, D. A., Narrow, W. E., Rae, D. S., Manderscheid, R. W., Locke, B. Z., & Goodwin, F. K. (1993b). The de facto US mental and addictive disorders service system. Epidemiologic catchment area prospective 1-year prevalence rates of disorders and services. *Arch Gen Psychiatry*, *50*(2), 85-94.
- Rollman, B. L., Hanusa, B. H., Gilbert, T., Lowe, H. J., Kapoor, W. N., & Schulberg, H. C. (2001). The electronic medical record. A randomized trial of its impact on primary care physicians' initial management of major depression [corrected]. *Arch Intern Med*, *161*(2), 189-197.
- Roman, P. M. (2002). Missing Work: The Decline in Infrastructure and Support for Workplace Alcohol Intervention in the United States, with Implications for Developments in Other Nations. In W. R. Miller & C. M. Weisner (Eds.), *Changing Substance Abuse Through health and Social Systems*. New York: Plenum Press.
- Rosenthal, M. B., Fernandopulle, H. R. S., & Landon, B. (2004). Paying for Quality: Provider' Incentives For Quality Improvement. *Health Aff (Millwood)*, *23*(2), 127-141.
- Rost, K., Humphrey, J., & Kelleher, K. (1994a). Physician management preferences and barriers to care for rural patients with depression. *Arch Fam Med*, *3*(5), 409-414.

- Rost, K., Smith, R., Matthews, D. B., & Guise, B. (1994b). The deliberate misdiagnosis of major depression in primary care. *Arch Fam Med*, 3(4), 333-337.
- Rushton, J. L., Felt, B. T., & Roberts, M. W. (2002). Coding of pediatric behavioral and mental disorders. *Pediatrics*, 110(1 Pt 1), e8.
- Saitz R, Mulvey KP, & Samet JH. (1997). The substance abusing patient and primary care: linkage via the addiction treatment system? *Substance Abuse*, 18, 187-195.
- Saitz, R., Mulvey, K. P., & Samet, J. H. (1997). The Substance Abusing Patient and Primary Care: Linkage via the Addiction Treatment System. *Substance Abuse*, 18, 187-195.
- Samet, J. H., Saitz, R., & Larson, M. J. (1996). A Case for Enhanced Linkage for Substance Abusers to Primary Medical Care. *Substance Abuse*, 17(4), 181-192.
- SAMHSA. (2001). *National Survey of Substance Abuse Treatment Services*.: Substance Abuse and Mental Health Services Administration.
- SAMHSA. (2004a). *Health Insurance Portability and Accountability Act: What it Means for Mental Health and Substance Abuse Services*, Substance Abuse and Mental Health Services Administration. November 4, 2004, from <http://hipaa.samhsa.gov/hipaacodes2.htm>
- SAMHSA. (2004b). *National Registry of Effective Programs and Practices (NREPP)*, Substance Abuse and Mental Health Services Administration. December 19, 2004, from <http://modelprograms.samhsa.gov/template.cfm?page=nrepbutton>
- SAMHSA. (2004c). *National Survey of Substance Abuse Treatment Services (N-SSATS), 2003*: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Office of Applied Studies.
- SAMHSA. (2004d). *Treatment Improvement Protocols*. Center for Substance Abuse Treatment (CSAT), Substance Abuse and Mental Health Services Administration (SAMHSA). Retrieved December 14, 2004, from <http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hstat5.part.22441>
- Saunders, R., & Heflinger, C. (2004). Integrating Data from Multiple Public Sources: Opportunities and Challenges for Evaluators. *Evaluation*, 10(3), 349-365.
- Saunders, R. C., & Heflinger, C. A. (2003). Access to and patterns of use of behavioral health services among children and adolescents in TennCare. *Psychiatr Serv*, 54(10), 1364-1371.
- Schizophrenia Patient Outcomes Team (PORT). (1998). *Treatment Recommendations*. Agency for Healthcare Quality and Research. Retrieved September 22, 2004, from <http://www.ahrq.gov/clinic/schzrec1.htm#info>

- Schulberg, H. C., Katon, W., Simon, G. E., & Rush, A. J. (1998). Treating major depression in primary care practice: an update of the Agency for Health Care Policy and Research Practice Guidelines. *Arch Gen Psychiatry*, 55(12), 1121-1127.
- Shapiro, H. (2004). Providers Must Participate in Creation of Performance Measures. *From the Field. Alcohol and Drug Abuse Weekly*, May 17.
- Shaul, J. A., Eisen, S. V., Stringfellow, V. L., Clarridge, B. R., Hermann, R. C., Nelson, D., et al. (2001). Use of consumer ratings for quality improvement in behavioral health insurance plans. *Jt Comm J Qual Improv*, 27(4), 216-229.
- Simpson, D. D., Joe, G. W., Rowan-Szal, G., & Greener, J. (1995). Client engagement and change during drug abuse treatment. *J Subst Abuse*, 7(1), 117-134.
- Smith, M. (2004, April 19-20). *Applying Community and Modularity Concepts to Consumer Surveys*. Paper presented at the 2nd Annual Forum on Performance Measurement in Behavioral Health and Related Service Systems, Atlanta, GA.
- Smith, M., Davis, S., Ganju, V., Tremaine, L., Adams, N., Felton, H., et al. (2004). Chapter 5. Twenty-five Years of the Mental Health Statistics Improvement Program: Past, Present, and Future- Where Have We Been and Where Are We Going? In R. Manderscheid & M. Henderson (Eds.), *Mental Health, United States, 2002* (Vol. DHHS Pub No. (SMA) 3938). Rockville, Maryland: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.
- Solberg, L. I., Mosser, G., & McDonald, S. (1997). The three faces of performance measurement: improvement, accountability, and research. *Jt Comm J Qual Improv*, 23(3), 135-147.
- Steinberg, E. P. (2003). Improving the Quality of Care -- Can We Practice What We Preach? *New England Journal of Medicine*, 348(26), 2681-2683.
- Strunk, B., & Hurley, R. (2004). *Paying for Quality: Health Plans Try Carrots Instead of Sticks*. Center for Studying Health System Change, Issue Brief No. 82. Retrieved, from www.hschange.org
- Sturm, R., & Sherbourne, C. D. (2001). Are barriers to mental health and substance abuse care still rising? *J Behav Health Serv Res*, 28(1), 81-88.
- Sturm, R., & Wells, K. B. (1995). How can care for depression become more cost-effective? *Jama*, 273(1), 51-58.

- Teague, G. B., Ganju, V., Hornik, J. A., Johnson, J. R., & McKinney, J. (1997). The MHSIP mental health report card. A consumer-oriented approach to monitoring the quality of mental health plans. Mental Health Statistics Improvement Program. *Eval Rev*, 21(3), 330-341.
- US Department of Health and Human Services. (1999). *Mental Health: A Report of the Surgeon general*. Rockville MD: US Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes of Medicine, National Institute of Mental Health.
- Veterans Health Administration. (2002). *FY2002 VHA performance measurement system. Technical Manual*. Office of Quality and Performance (10Q). Retrieved December 16, 2004, from www.oqp.med.va.gov/cpg/MDD/P/MDD_3_8_02_techman.doc
- Wagner, E. H., Austin, B. T., & Von Korff, M. (1996). Organizing care for patients with chronic illness. *Milbank Q*, 74(4), 511-544.
- Walkup, J. T., Boyer, C. A., & Kellermann, S. L. (2000). Reliability of Medicaid claims files for use in psychiatric diagnoses and service delivery. *Adm Policy Ment Health*, 27(3), 129-139.
- Wang, P. S., Berglund, P., & Kessler, R. C. (2000). Recent care of common mental disorders in the United States : prevalence and conformance with evidence-based recommendations. *J Gen Intern Med*, 15(5), 284-292.
- Wang, P. S., Demler, O., & Kessler, R. C. (2002). Adequacy of treatment for serious mental illness in the United States. *Am J Public Health*, 92(1), 92-98.
- Weisner, C., & Matzger, H. (2002). A prospective study of the factors influencing entry to alcohol and drug treatment. *J Behav Health Serv Res*, 29(2), 126-137.
- Weisner, C., Matzger, H., Tam, T., & Schmidt, L. (2002). Who goes to alcohol and drug treatment? Understanding utilization within the context of insurance. *J Stud Alcohol*, 63(6), 673-682.
- Weisner, C., & Schmidt, L. (1993). Alcohol and drug problems among diverse health and social service populations. *Am J Public Health*, 83(6), 824-829.
- Weisner, C., & Schmidt, L. A. (2001). Rethinking access to alcohol treatment. In M. Galanter (Ed.), *Recent Developments in Alcoholism, Volume 15. Services Research in the Era of Managed Care: Organization, Access, Economics, Outcome* (pp. 107-136). New York: Kluwer Academic/Plenum Publishers.
- Wells, K., Klap, R., Koike, A., & Sherbourne, C. (2001). Ethnic disparities in unmet need for alcoholism, drug abuse, and mental health care. *Am J Psychiatry*, 158(12), 2027-2032.

- Whitlock, E. P., Polen, M. R., Green, C. A., Orleans, T., & Klein, J. (2004). Behavioral counseling interventions in primary care to reduce risky/harmful alcohol use by adults: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med*, *140*(7), 557-568.
- Wickizer, T., Maynard, C., Atherly, A., Frederick, M., Koepsell, T., Krupski, A., et al. (1994). Completion rates of clients discharged from drug and alcohol treatment programs in Washington State. *Am J Public Health*, *84*(2), 215-221.
- Wickizer, T. M., Campbell, K., Krupski, A., & Stark, K. (2000). Employment outcomes among AFDC recipients treated for substance abuse in Washington State. *Milbank Q*, *78*(4), 585-608, iv.
- Wilk, A. I., Jensen, N. M., & Havighurst, T. C. (1997). Meta-analysis of randomized control trials addressing brief interventions in heavy alcohol drinkers. *J Gen Intern Med*, *12*(5), 274-283.
- Wingert, T. D., Kralewski, J. E., Lindquist, T. J., & Knutson, D. J. (1995). Constructing episodes of care from encounter and claims data: some methodological issues. *Inquiry*, *32*(4), 430-443.
- Wisconsin Department of Health and Family Services. (2003). *MEDDIC-MS. Medicaid encounter data driven improvement core measure set. Madison (WI)*. Division of Health Care Financing, Bureau of Managed Health Care Programs, State of Wisconsin. Retrieved, from <http://dhfs.wisconsin.gov/medicaid7/index.htm>
- Wolff, N. (2001). Randomized trials of socially complex interventions: promise or peril? *J Health Serv Res Policy*, *6*(2), 123-126.
- World Health Organization Regional Office for Europe. (2004). *What is the effectiveness of old-age mental health services?* January 5, 2005, from http://www.euro.who.int/HEN/Syntheses/mentalservice/20040721_6
- Wynia, M. K., Cummins, D. S., VanGeest, J. B., & Wilson, I. B. (2000). Physician manipulation of reimbursement rules for patients: between a rock and a hard place. *Jama*, *283*(14), 1858-1865.